

Lab 1: Gaia, RR Lyrae Stars, and Galactic Dust

Andrew Goh

Oct 2022

1 Introduction

Cosmic dust permeates our universe, originating from many astrophysical collisions, comets, supernova remnants, and events dating all the way back to the Big Bang. Although the dust grains and gas are small, hence the name ‘dust’, their presence can have a significant impact on astrophysical measurements. Extinction is the phenomenon in which light from a star or a bright source, becomes scattered or absorbed while traveling along the line of sight due to cosmic dust between us and the source. Shorter wavelengths of light have a higher chance of being scattered or absorbed by dust, while longer wavelengths tend to pass through dust. From our perspective, this will result in us seeing the source more red than its original (intrinsic) color, and is thus called ‘interstellar reddening’.

Consequentially, extinction must be taken to account when making photometric measurements making it crucial to various astrophysical studies. The most widely used extinction map is the SFD map by Schlegel, Finkbeiner, and Davis 1998 [3]. We attempt to construct a crude but reasonable extinction map of the sky, and make comparisons to this SFD map.

2 Methods

We utilize the concept of a “standard candle”, which is a specific astronomical object that has a known intrinsic brightness or absolute magnitude. By identifying the object in the sky and measuring its apparent magnitude, we can determine the distance to the object given our knowledge of its absolute magnitude. In this lab, we make use of the Leavitt Law, which defines the relationship between the luminosity and the period of a pulsating star, known as an RR Lyrae star. An RR Lyrae star has a variable luminosity, in which its period can range from a few hours to two days. Utilizing the Leavitt Law, if we know the pulsation period of the star, then we can determine its intrinsic color and absolute magnitudes. With this, we can calculate the color excess:

$$E(G_{BP} - G_{RP}) = (G_{BP} - G_{RP})_{observed} - (G_{BP} - G_{RP})_{intrinsic} \quad (1)$$

With interstellar reddening, we can expect the observed bp-rp color to be greater than the intrinsic color, giving us a positive color excess value. By querying the `vari_rrlyrae` catalog and cross matching with the Gaia EDR3 source catalog from the Gaia archive, we can calculate the color excess and G-band extinction for each star, and plot the values accordingly to each star’s galactic coordinates, thus effectively creating a dust map of our observed sky.

As a means of demonstrating our understanding as well as justifying our procedures, we derive and calculate the magnitude and periods from raw light curves of RR Lyrae stars using

Fourier analysis. We then employ Markov chain Monte Carlo samplers to derive a linear model for the period-luminosity relationship given the Gaia data after verifying our calculations of magnitudes and periods. We use this linear model to inform us of the intrinsic bp-rp color of a specific RR Lyrae star.

3 Fourier Analysis of Light Curves

In this section, we demonstrate the periodicity of RR Lyrae pulsations and estimate their mean magnitudes by applying a Fourier decomposition on the data. Doing this circumvents the inconsistent intervals in which sources are measured, and provides us with a model light curve function, which then can be used to calculate a reasonable mean magnitude. We use ADQL to query the selection of data we want from the Gaia database. We can query our selection directly from the Jupyter Notebooks and save the data through importing Gaia from the astroquery package. The raw light curves are accessed via the Datalink protocol given the source ID's of our query. We are interested in the epoch photometry tables. Figure 1 shows a plot of the epoch photometry of an arbitrary source we have chosen with a high signal to noise ratio.

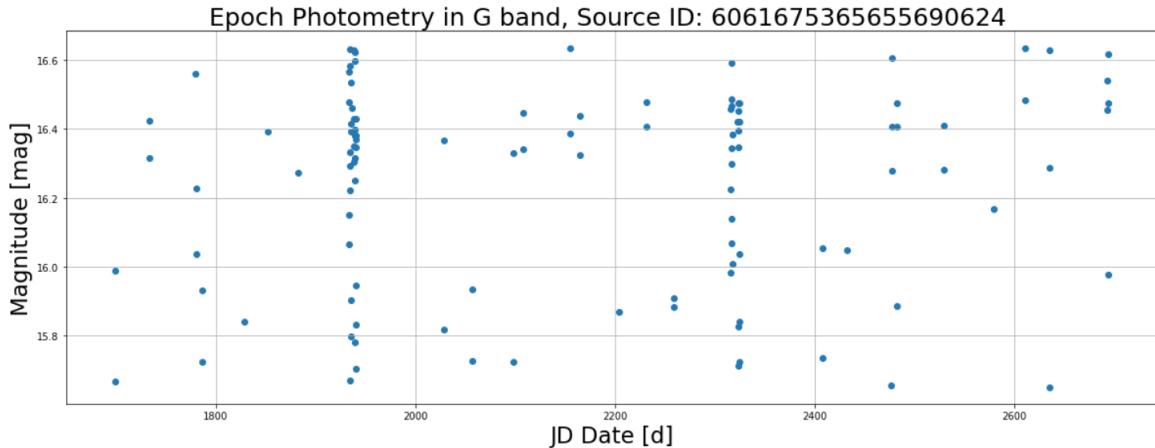


Figure 1: Epoch photometry of an arbitrary RR Lyrae source (Gaia source ID is labeled in the title) plotting measured magnitude versus Julian Date. We can see that the source is measured at inconsistent intervals, but there is a clear indication of consistent variability in its magnitude.

3.1 Lomb Scargle Periodogram

The motive of using the Lomb Scargle Periodogram is to Fourier Transform the epoch photometry signal in order to find the period. Since the data is sampled at irregular intervals with measurement errors, we can only approximate a model as a sum of sine curves over a frequency range:

$$y_{model}(t; f) = y_0(f) + A_f \sin(2\pi f(t - \phi_f)) \quad (2)$$

which the parameters of the model, A_f , ϕ_f , $y_0(f)$ are fitted to minimize χ^2

$$\chi^2(f) = \sum \left(\frac{y(t) - y_{model}(t; f)}{\sigma(t)} \right)^2 \quad (3)$$

where $y(t)$ is the measured magnitude at a specific Julian Date (JD) and $\sigma(t)$ is the measurement error at $y(t)$. Using the `astropy.timeseries.LombScargle` package, we can create a Lomb Scargle periodogram which is a plot of the power spectrum A_f versus frequency f . Figure 2 shows a representation of the periodogram. We assume that the frequency that contains the highest A_f value is the dominant frequency, corresponding to the fundamental pulsation mode of the RR Lyrae star. We obtain a power spectrum and the period of a sample of 100 RR Lyrae stars that have a measured period at the fundamental pulsation mode, corresponding to the ADQL syntax `pf IS NOT NULL`. This query returns the class of RRab stars, which their pulsations occur at the fundamental mode. This effectively filters out the class of RRc stars, which are first overtone pulsators. This specific class of RR Lyrae stars have no measured fundamental mode periods, since it undergoes characteristic modulations in frequency. We also require that at least 40 clean epochs have been measured: `gaiadr3.vari_rrlyrae.num_clean_epochs_g >= 40`. Figure 2 also showcases our comparisons of Lomb Scargle obtained periods with reported Gaia periods. Most of the sources have close matching correspondences, but there are some observed small deviations and two large deviations. These deviations could likely be due to the Blazhko effect [2], the phenomenon in which some RR Lyrae stars can exhibit a characteristic modulation in phase and amplitude. This phenomenon occurring in conjunction with irregularly measured intervals can contribute to an inaccurate calculation of a power spectrum.

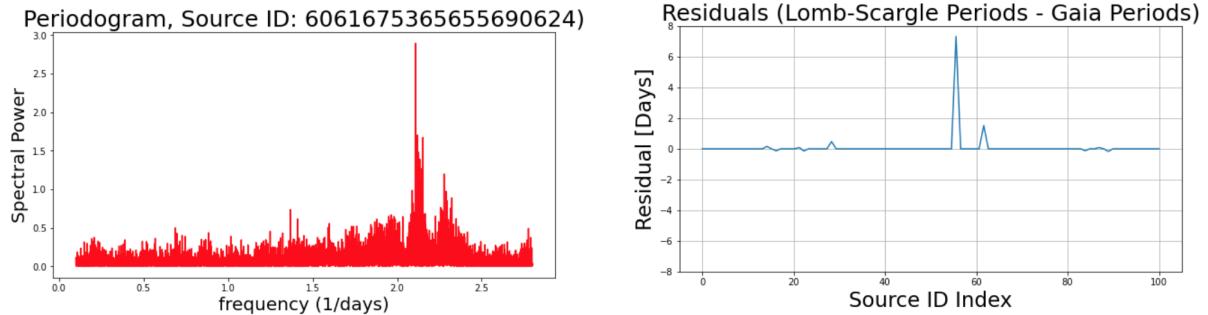


Figure 2: The right figure shows a Lomb Scargle Periodogram for the same source that has a high s/n ratio. We can see a high power amplitude peak that corresponds to the frequency of the fundamental pulsation mode compared to the low fuzzy noise levels. The left figure shows a residual plot between 100 periods obtained from the Lomb Scargle periodograms and the periods reported in the `vari_rrlyrae` catalog. Deviations from the zero are likely due to the Blazhko effect or either to large measurement noise, hampering our method of obtaining a significant peak of the fundamental mode period.

3.2 Fourier series representation of light curves

Since the magnitudes of the raw light curves are measured at irregular intervals, simply taking the average of these measurements is practically incorrect if we want to calculate mean magnitudes. A more robust way of calculating mean magnitudes is to represent our data as a sum of sine and cosines in a similar fashion to the Lomb Scargle analysis. We define our light curve model to be:

$$f(t) = A_0 + \sum_{k=1}^K [a_k \sin(k\omega t) + b_k \cos(k\omega t)] \quad (4)$$

where A_0 , a_k , b_k are parameters to be fitted when minimizing the residual between the model and the measured Gaia data points, and ω is the frequency simply derived from the inverse of the Lomb Scargle period. As K increases, we expect that it represents our data with more

harmonics, resulting in a more accurate model fit. In order to solve for the parameters, we can utilize a simple linear optimization method by recasting the problem in matrix form:

$$\mathbf{X}b = y \quad (5)$$

$$\mathbf{X} = \begin{bmatrix} 1 & \sin(wt_1) & \cos(wt_1) & \sin(2wt_1) & \cos(2wt_1) & \sin(3wt_1) & \cos(3wt_1) & \dots \\ 1 & \sin(wt_2) & \cos(wt_2) & \sin(2wt_2) & \cos(2wt_2) & \sin(3wt_2) & \cos(3wt_2) & \dots \\ \vdots & \vdots \end{bmatrix} \quad (6)$$

$$b = \begin{bmatrix} A_0 \\ a_1 \\ b_1 \\ a_2 \\ b_2 \\ a_3 \\ b_4 \\ \vdots \end{bmatrix} \quad (7)$$

$$y = \begin{bmatrix} y(t_1) \\ y(t_2) \\ \vdots \end{bmatrix} \quad (8)$$

The \mathbf{X} has $2K+1$ columns, depending on how many harmonics we decide to use, and the number of rows is the number of measured points the light curve has given its time and magnitude. The b vector has $2K+1$ elements, containing all of the fitting parameters. The y vector contains the measured magnitudes at a given time. We “solve” equation (5) with the `np.linalg` package using the least squares method `lstsq`. This method will return the parameters of best fit given a certain K . We can then reconstruct an “optimally measured” light curve with the model by setting an arbitrary range for t , and simply carrying out the multiplication $\mathbf{X}b$. In order to better visualize our fitted model, we plot the epoch photometry in which each data points are set at its corresponding phase according to the Lomb Scargle period, such that all of the data points are seen within one cycle. This is done by taking the measured time modulo the period. Figure 3 shows a plot of the phased epoch photometry light curve along with the fitted Fourier series representations (FSR) at different values of K .

3.3 Cross-validation

To gauge the efficacy of this fitting method at different K ’s, we use a cross-validation technique common in machine learning. We first randomize and split our measured data into training data and testing data, where we designate 80% of the measured data to the training data and 20% to the testing data. We do this using the `sklearn.model_selection` package. We fit our model, equation (5), only to the training data set. We then compute χ^2/N_{data} for the fitted model against the training data set, and then against the testing data set. We do this for K ranging from 1 to 25, which we found an optimal K to be 7. The problem of going to higher K values is that the least squares algorithm simply minimizes differences between the model and the individually sampled data points. It does not take any inference in a curve that the data

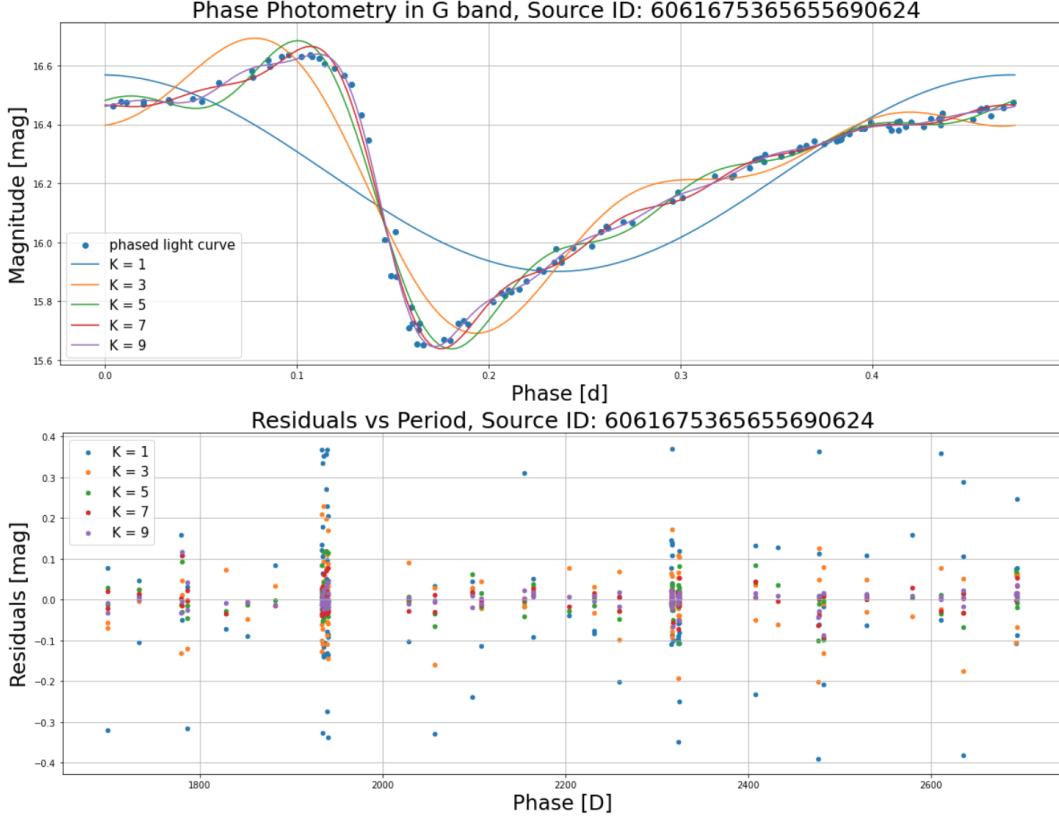


Figure 3: The top figure is the phased photometry of the same source as used above plotted against the fitted FSR model. By plotting the phased data points, we can see a general curve that is implied by the data points. As K increases, the fitted model becomes more representative the general curve. The bottom figure shows the residuals between the fitted model at each K and the measured magnitudes.

points may draw out. Thus, the model fit may be a curve that is not actually representative of the curve implied by the data points, but just so happens to pass close to the data points as it is simply required to by the least squares method. Figure 4 shows a plot of χ^2/N_{data} against a range of K for an arbitrary source.

3.4 Light Curve Extrapolation and mean magnitudes

By choosing the optimal $K = 7$, we can then use this value to do the rest of our light curve fits. We can extrapolate the fitted model over future days. This effectively predicts future magnitudes. Figure 5 shows an extrapolated light curve plot.

Since we can reconstruct reasonable light curves, we can now estimate an accurate mean apparent magnitude. For the same 100 sample of RR Lyrae stars, we calculate their mean magnitudes using the FSR model. However, since magnitudes are logarithms, we must convert the logarithms to fluxes, take the average, and convert the average back in to logarithms. The procedure for adding the magnitudes is as follows:

$$m = -2.5 \log_{10}(10^{-m_1/2.5} + 10^{-m_2/2.5} \dots) \quad (9)$$

We also plot the residuals with the reported magnitudes to compare how well the models can give us an accurate mean magnitude.

3.5 A comparison of RRab and RRc light curves

With the FSR models, we can plot light curves of both classes of RRc and RRab stars and visually observe their differences. As shown in Figure 7, the RRc stars exhibit a more sinusoidal shape and one is shown to have noticeable intrinsic scatter, indicating changes in its pulsation period. The RRab class of RR Lyrae stars have a characteristic sharp dip in magnitude in which it quickly rises again. In both the RRab and RRc class our FSR fits do not capture the deviations of data points from the fitted curve, which may be indications of phase or amplitude modulation. Therefore we must take our model with a sense of caution.

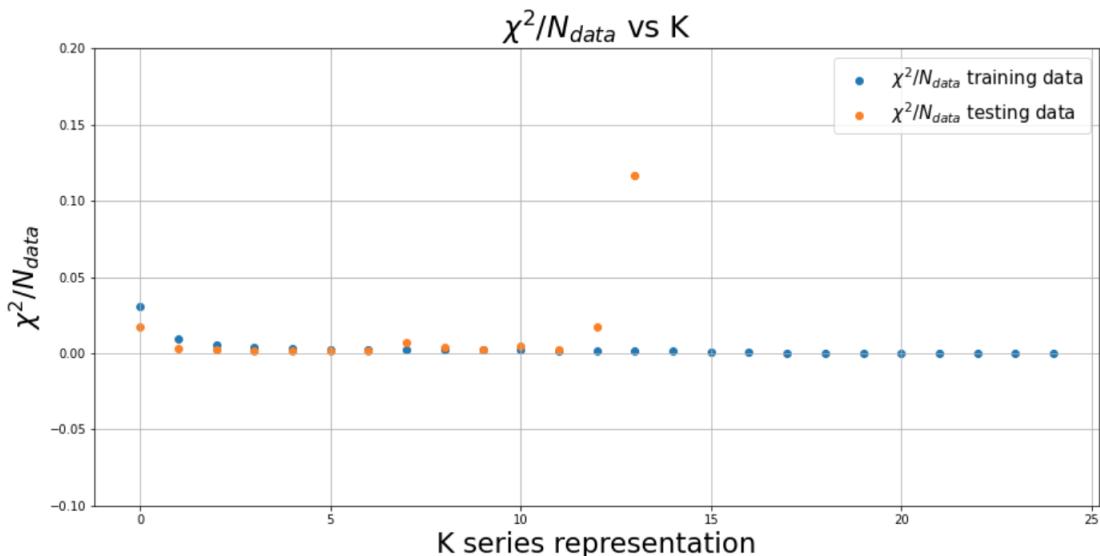


Figure 4: Plot of χ^2/N_{data} versus K for an arbitrary source. We can see that at larger K's, the fitting method is unstable and the discrepancy diverges, indicating that the algorithm is finding a curve that passes through the training data set points but is not actually representative of the correct curve, as the training data χ^2/N_{data} values indicate. The values at which this divergence happens occurs between different sources.

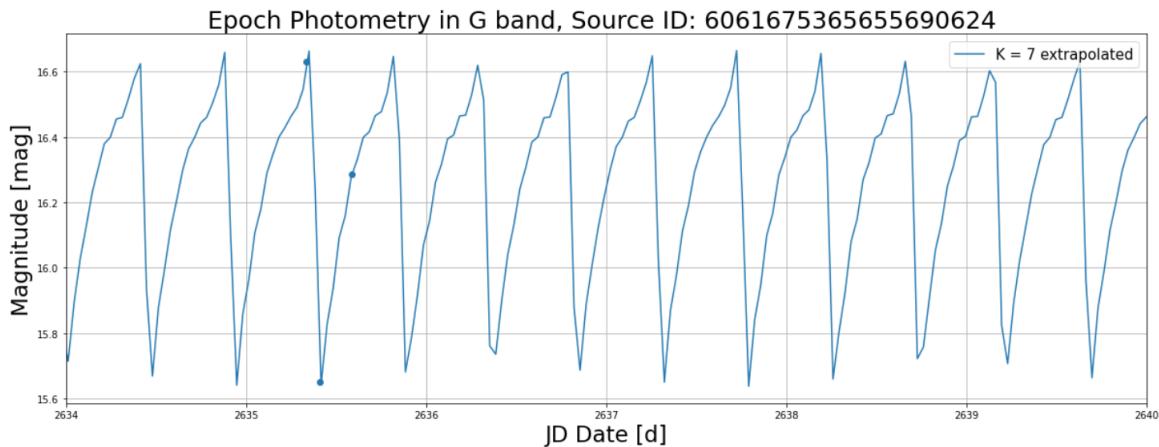


Figure 5: Extrapolated light curve plot several days after the latest measured magnitude reported in the Gaia archive.

	source_ids	fourier_mean_mag
0	245002531050576896	15.857052
1	245504251951140864	17.990431
2	245823861938360064	19.337150
3	246756973652292992	12.355293
4	359112665277840512	15.905017
...
95	2195648132814029312	19.068327
96	2195750146876812032	19.203695
97	2195878209919982464	17.621507
98	2195880752540525952	16.563064
99	2195937888486787840	18.360417

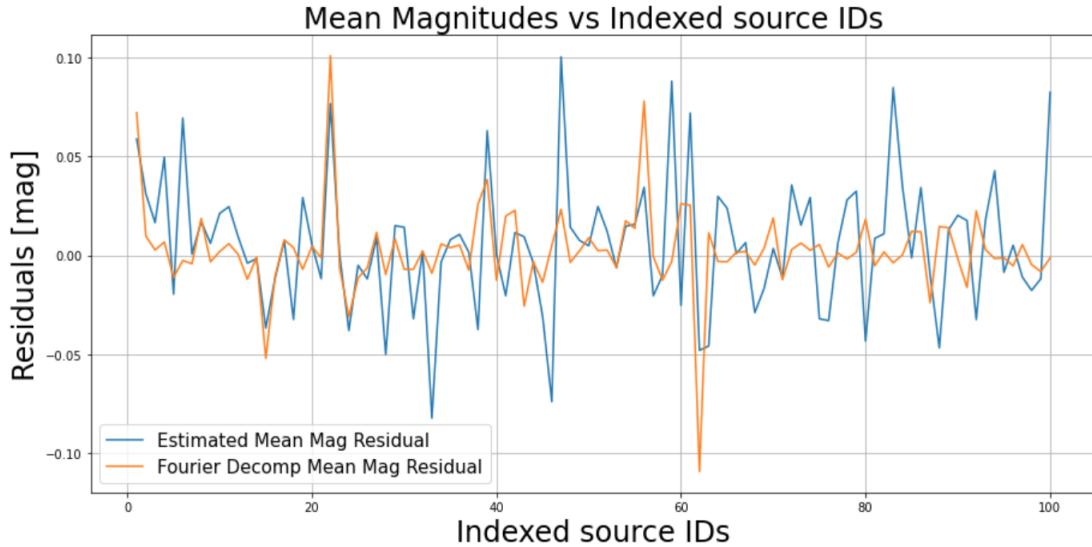


Figure 6: The top figure shows a table listing mean magnitudes computed through the Fourier series representation (FSR) by Gaia source IDs. The plot shows the residuals between the reported Gaia magnitude and the mean magnitudes obtained by the FSR and the residuals of mean magnitudes obtained by naively calculating the average of the measured data in the raw light curves. We can observe that for most sources, the residuals from the FSR are smaller than the naive averaging of measured data. However there are similar instances of deviation between the FSR and naive averaging, which may again be a consequence of the Blazhko effect.

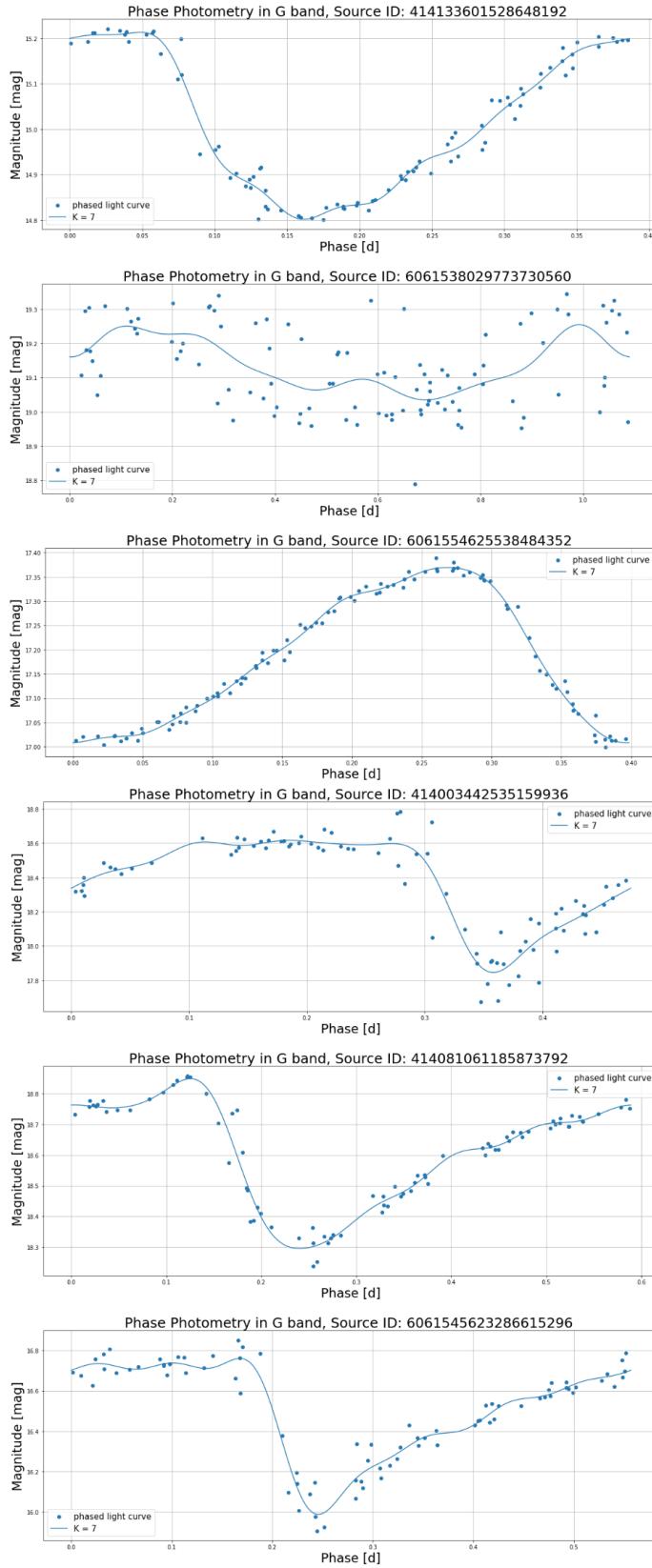


Figure 7: Light curves with FSR model over plotted. The top 3 plots are of RRc class stars and the bottom 3 are of RRab class stars. The RRc stars have a clear sinusoidal shape. The middle RRc plot has a clear intrinsic scatter, where we can assume this is due the characteristic of RRc stars having a modulating frequency.

4 Absolute Magnitudes and Distances

After demonstrating that mean magnitudes and periods of RR_{ab} stars reported in the Gaia tables are reliable we now move on to calculating the absolute magnitudes. Given an apparent mean magnitude, we can calculate absolute magnitude given the source's distance:

$$M = m - 5 * \log_{10}(d) + 5 \quad (10)$$

Where M is the absolute magnitude, m is the apparent magnitude, and d is the distance to the source in parsecs. We query 500 RR Lyrae stars outside the galactic disk ($|b| > 30$ degrees) and less than 4kpc away to account for minimal extinction, as we assume most of the dust is in the galactic disk. The `vari_rrlyrae` catalog does not contain parallax measurements so we must cross match and join with the `gaia_source` catalogs in our ADQL query.

4.1 Bailer-Jones Parallaxes

The relationship between parallax and distance requires a non linear transformation. As parallax measurements become smaller, the errors on distances become significantly large. Using Gaia reported parallaxes can hamper our calculations for absolute magnitude, so we make use of constraints on distances using the Bailer Jones distance estimates [1] which uses Bayes' theorem. Bayes' theorem outlines a proportionality between a posterior probability distribution, which is the probability of a value given a calculated model, the likelihood probability distribution, which is the probability of the model given our value, and the prior, which is a probability distribution consisting of our preconceived knowledge. An unnormalized equation of Bayes' theorem is presented as:

$$P(r|\theta) \propto P(\theta|r)P(\theta) \quad (11)$$

Where $P(r|\theta)$ is the posterior probability distribution, $P(\theta|r)$ is the likelihood probability distribution, and $P(\theta)$ is the prior. The posterior probability distribution for distance is calculated by using a prior based on the knowledge of our galactic structure. This distance prior is an exponentially decreasing function with a characteristic length scale that is direction dependent. The likelihood is an assumed Gaussian that depends on the parallax zero point. The posterior distribution is calculated numerically using Monte Carlo Markov Chain samplers. We can query the Bailer Jones distance estimates using the `external.gaiaedr3_distance` external catalog. The columns we are interested in are `r_hi_geo` - `r_lo_geo`, which represents the 68% confidence interval, and `r_med_geo` which is the median of the posterior distribution, all of which are in units of parsecs. We take this median to be the distances of interest over the parallax measurements.

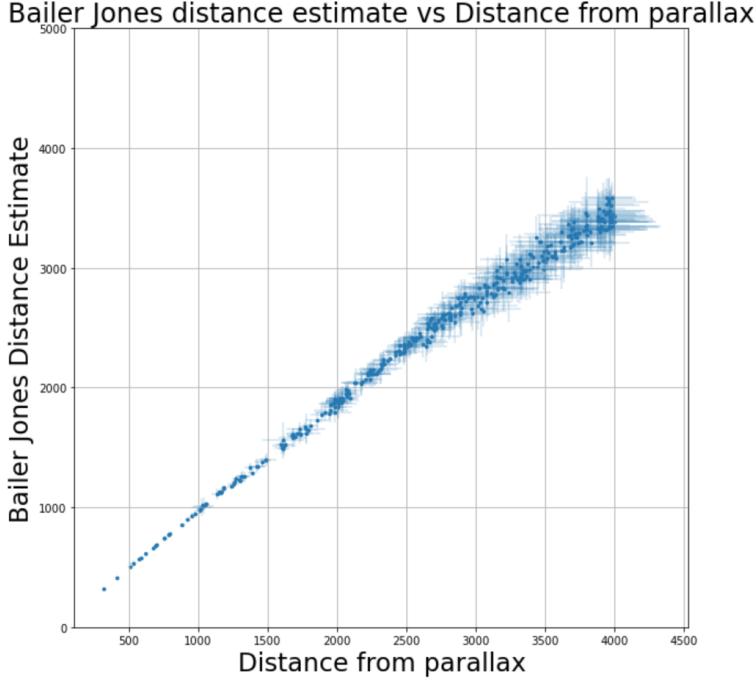


Figure 8: Plot of Bailer-Jones distance estimate versus distance from parallax (where distance from parallaxes is calculated as $1/\varpi$. As distances become larger, the discrepancies between the Bailer-Jones distances and the parallax become larger, as well as the error bars calculated from `r_hi_geo - r_lo_geo` and parallax errors reported in the `gaia_source` catalog.

5 Period-Absolute Magnitude relations

Using the Bailer-Jones distances, we use equation (10) to calculate the absolute magnitude for each of the 500 RR Lyrae stars we queried, where we substitute d with `r_med_geo`. We can then plot these magnitudes against the Lomb Scargle periods. Upon plotting, we can see that there are deviations from the cluster that are due to inaccurate measurements in parallax. We make a crude cut to these sources by simply applying a cut on absolute magnitude within the ADQL query: `WHERE vari_rrlyrae.int_average_g - 5 * LOG10(gaiaedr3_distance.r_med_geo) + 5 < 3.5`.

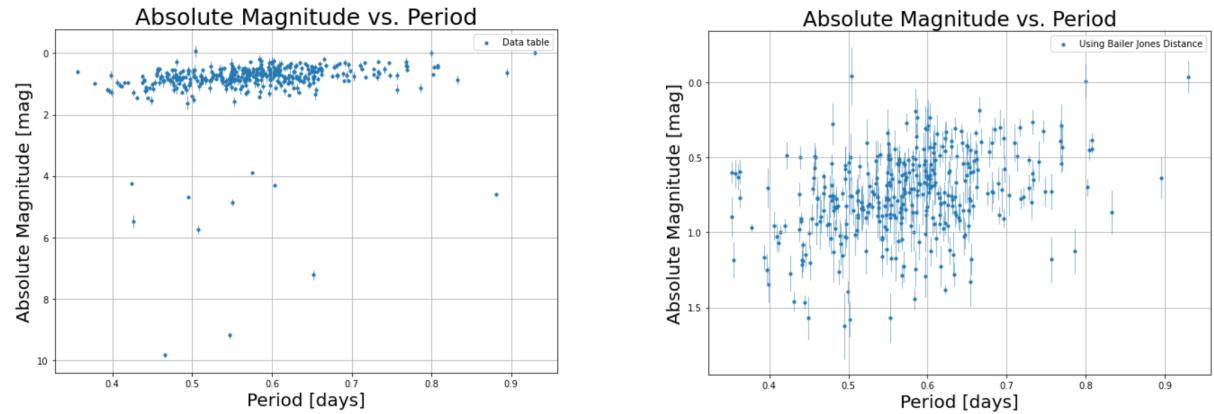


Figure 9: Plot of Absolute Magnitude versus Period of 500 RR Lyrae stars. The left plot shows the uncut data set. We can see a clear cluster indicating a positive correlation between absolute magnitude and period. However, there are sources that seem to deviate far from this cluster due to large parallax measurement errors. The right plot shows the data set after applying cuts on absolute magnitude below an estimated threshold of 3.5 mag.

6 MCMC methods for fitting Period-Luminosity relations

Markov Chain Monte Carlo (MCMC) methods are an effective way to not only fit a model to a data set, but also to provide posterior distributions to the parameters of the model. In other words, it can tell us how certain our fitted parameters are and to what extent that they are correlated. In the previous section, we have obtained absolute magnitudes and periods which showcase a positive correlation, however there is a significant spread to this data. Using MCMC, we can fit a line to this data, and also effectively fit its intrinsic spread.

6.1 Metropolis-Hastings Algorithm

In order to develop a sufficient understanding of MCMC methods, we hard code the so-called Metropolis-Hastings algorithm. The logic is outlined below

- Define our data:
 - P_i = array of periods
 - M_i = array of absolute magnitudes
 - σ_i = array of measured error of absolute magnitude
- Define the model: $M_G(a, b, P) = a \times \log_{10}(P/day) + b$. This assumes for our model that magnitude and period be linearly correlated in log space. The parameters a and b are to be fitted.
- Define the likelihood distribution and take its natural log:

$$\mathcal{L} = \prod_i \frac{1}{\sqrt{2\pi(\sigma_i + \sigma_{scatter})^2}} e^{\frac{-(M_i - M_G(a, b, P_i))^2}{2(\sigma_i + \sigma_{scatter})^2}} \quad (12)$$

This assumes a normal distribution centered at the measured absolute magnitude, where the larger the residual between the measured magnitude and our model, the smaller the probability or likelihood. We are also fitting for $\sigma_{scatter}$, which this is the intrinsic scatter of our data that is apart from measurement error. The total likelihood is a product of these probabilities. We take the natural log of the likelihood because it is more numerically stable to take into computation, as the exponential component can output very large or small values that our program can not handle.

- Initialize starting points for the parameters to be fitted and set these as the ‘current’ parameters.
- Generate proposal parameters using a random walk method, i.e. drawing a sample from a narrow gaussian distribution centered on the current parameters. We can manually scale the step-size of our walk such that it explores a reasonable parameter space.
- Calculate the likelihood distribution for both the current parameters and the proposed parameters.
- Take the ratio of these likelihood distributions, also known as the ‘acceptance fraction’ $\mathcal{L}_{proposed}/\mathcal{L}_{current}$

- Decide whether to accept or reject the proposal parameters
 - If the proposed likelihood is greater than the current likelihood, we will always accept the proposal parameters.
 - If the proposal likelihood is less than the current likelihood, then the chances of accepting the proposal is the acceptance fraction.
- For each iteration, save the parameters into a list, called the trace.
- Plotting the traces as a histogram will showcase the effective posterior distribution of the parameters.
- The step size must be fine tuned such that the acceptance fraction is within the order of 0.5 to generate reasonable traces and histograms.

We use this algorithm to fit the absolute magnitude versus period sample of RR Lyrae stars in Figure 9. Below are figures showcasing the results of the Metropolis Hastings algorithm.

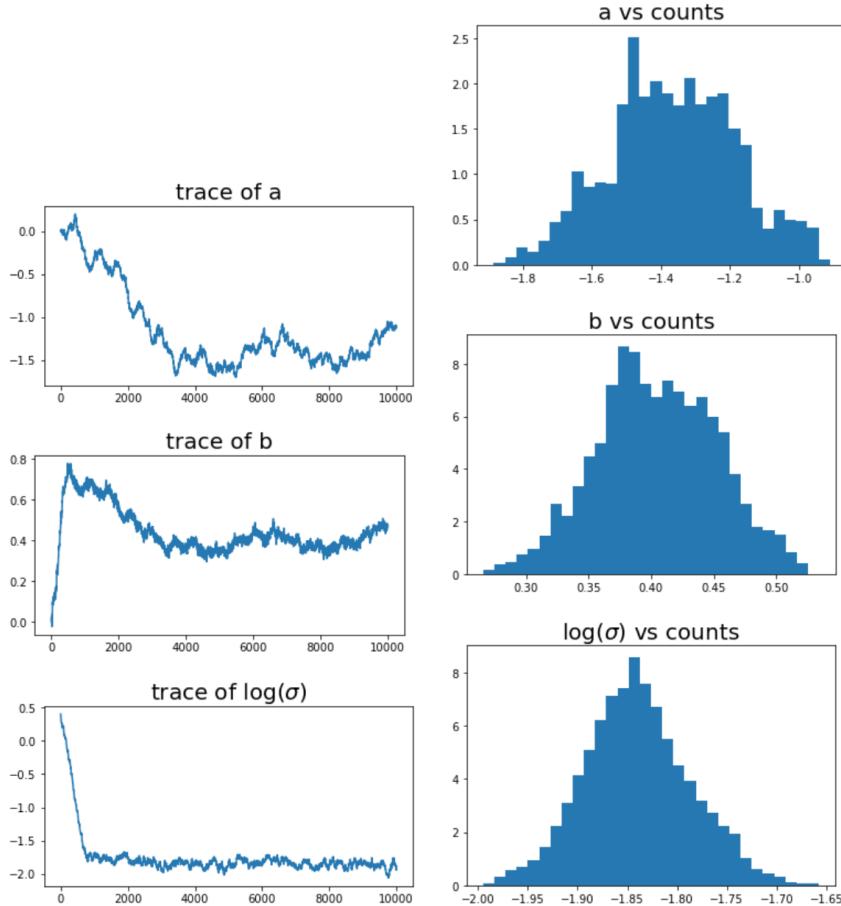


Figure 10: The left figures are plots of the traces, showcasing the random walks that are determined by the acceptance fraction. We can see that after there is a convergence on the parameters, however somewhat weak and prone to deviation. The right figures are histograms of the trace, representing a posterior distribution of the parameters. The standard deviation and mean can be obtained from these histograms.

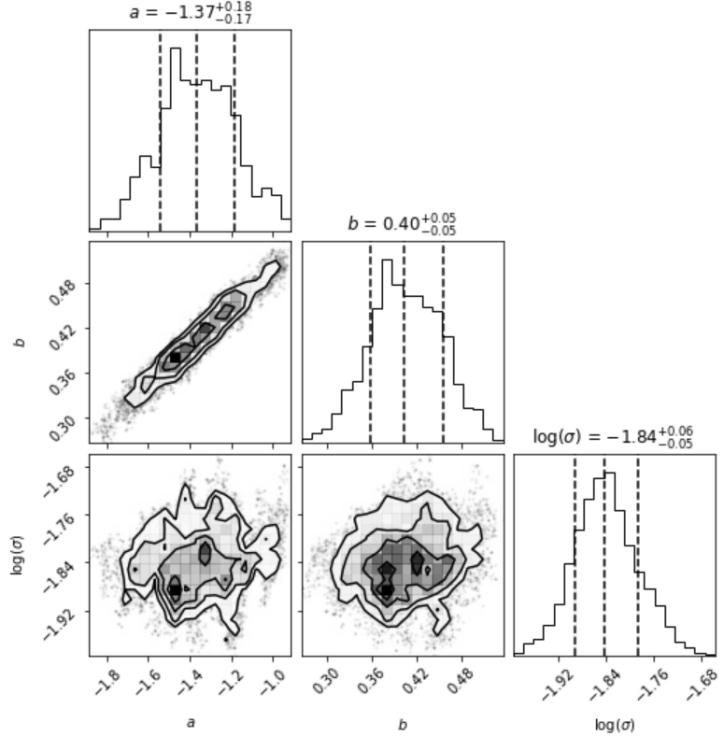


Figure 11: A corner plot that plots the individual distributions on the diagonals and the 2D confidence intervals on the lower diagonals. As expected, the a and b parameters to be correlated as the slope and magnitude of a line can both be adjusted in correspondence for a best fit line.

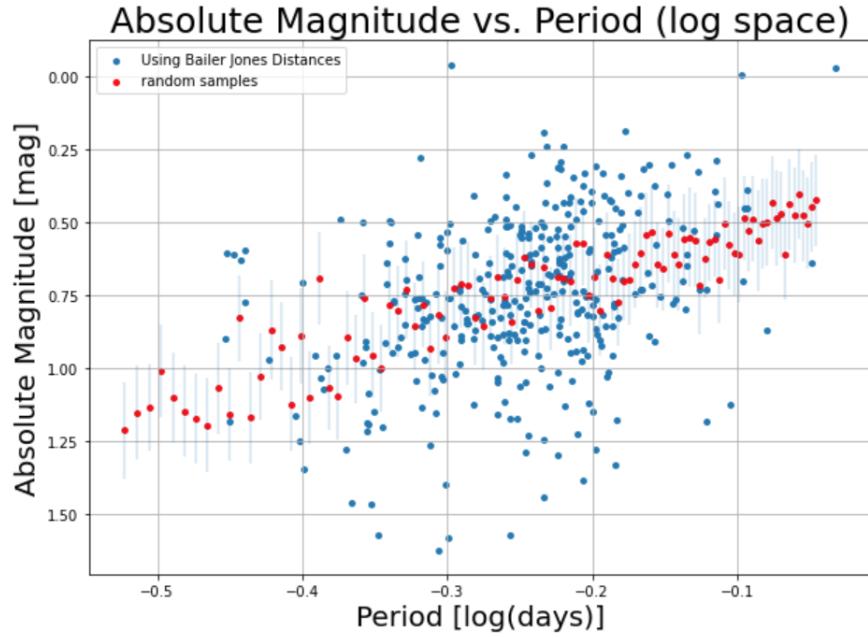


Figure 12: Absolute Magnitude vs Period data set overplotted with random samples drawn from the parameters' posterior distributions. We can see that the random samples represents the linear trend fairly reasonably, and the error bars, which are the fitted errors, appear to cover a standard deviation (68%) of the spread of data.

6.1.1 PYMC3

We then turn to the no-U-turn Hamiltonian Monte Carlo sampler which is a more efficient and robust algorithm. The algorithm utilizes Hamiltonian mechanics which generates ‘momentum’ for the parameter search to converge quickly. The NUTS sampler is able automatically optimize the step size of the searches, which is superior in comparison to the Metropolis-Hastings algorithm where the step size needs to be fine tuned to achieve reasonable results. We can use the NUTS sampler provided in the `pymc3` package.

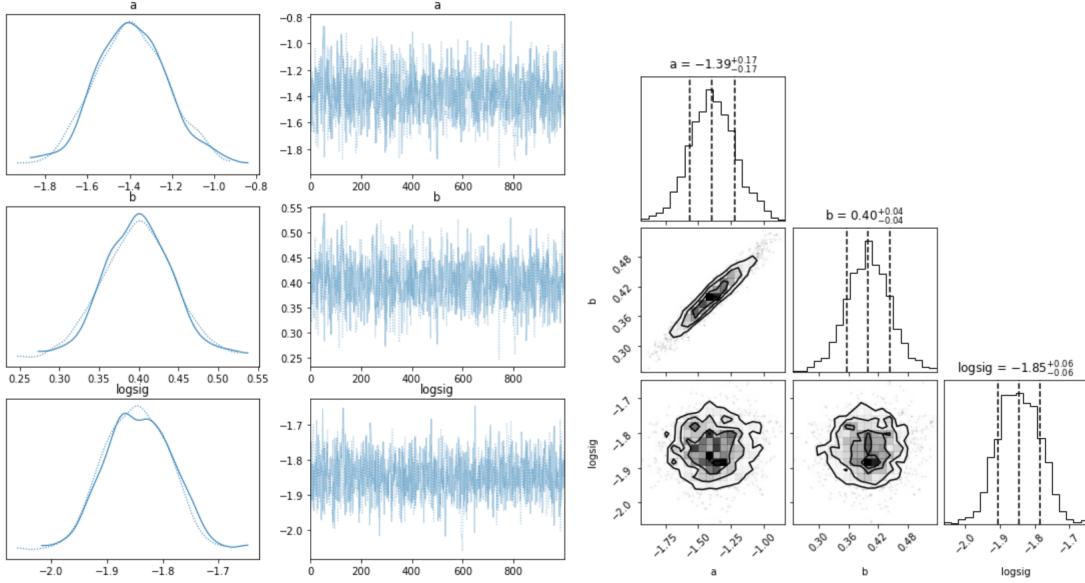


Figure 13: Trace and corner plots of the NUTS sampler in the `pymc3` package. The observed distributions here are much more uniform and clean as compared with the results of the Metropolis Hastings algorithm, however, both methods are able to converge to the same mean and standard deviations of the parameters.

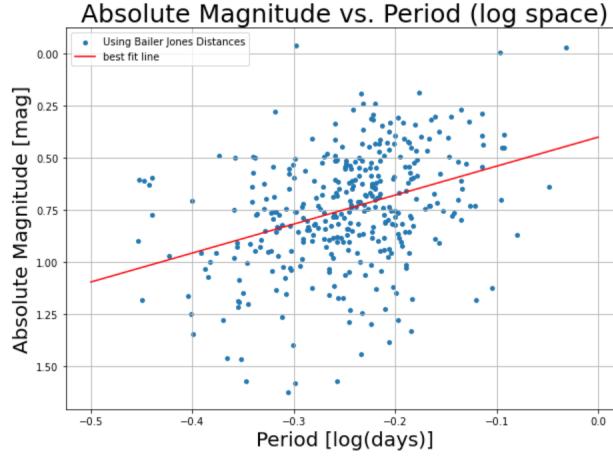


Figure 14: Best fit line using the mean of the posterior distributions of the parameters obtained from the NUTS sampler plotted against the same data as above.

6.2 Applying MCMC to sources in the WISE survey

We can now observe period luminosity relations in different bands by utilizing the NUTS sampler. The WISE catalog allows us to fit period luminosity relations for the W2 magnitude. We cross match our data set of 500 RR Lyrae stars with the external WISE catalog `gaiadr1.allwise_original_valid`.

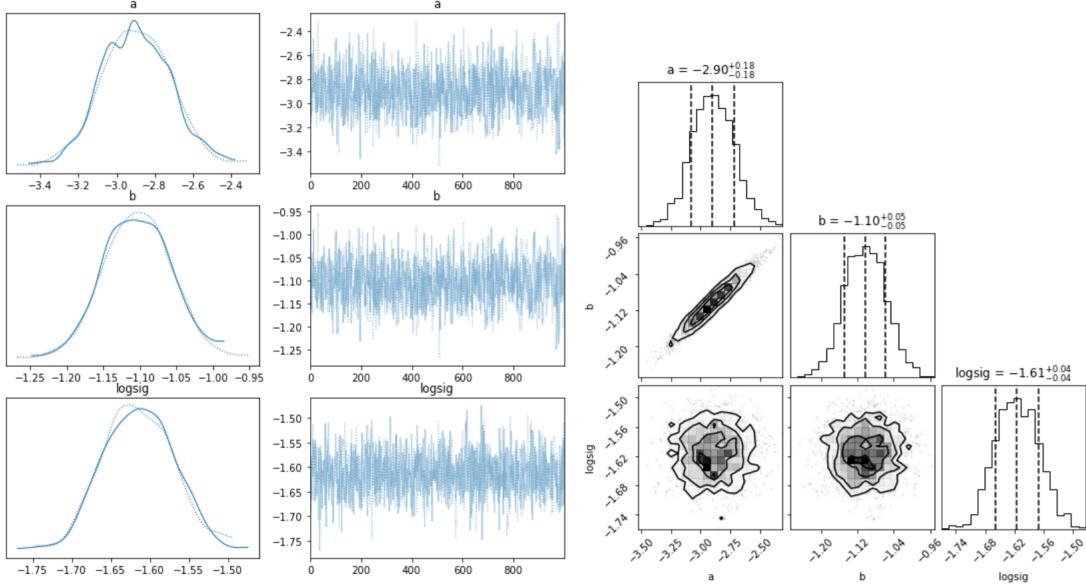


Figure 15: Trace and corner plots of the NUTS sampler fitting WISE w2mpro magnitudes versus period.

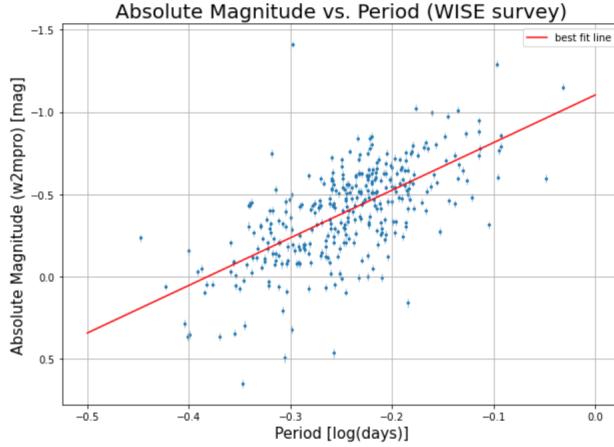


Figure 16: Best fit line using the mean of the posterior distributions of the parameters obtained from the NUTS sampler plotted against the same data as above.

We find that the period luminosity relationship is steeper in the W2 band compared to the optical G band.

6.3 Applying MCMC to derive color-period relation

Since our data set consists of RR Lyrae close to us and outside the galactic disk, we can expect that there is little to no extinction that affects to the bp-rp color. We can thus derive a model for the intrinsic bp-rp color of an RR Lyrae star given its period. Again, we are fitting a line to the data in \log_{10} space.

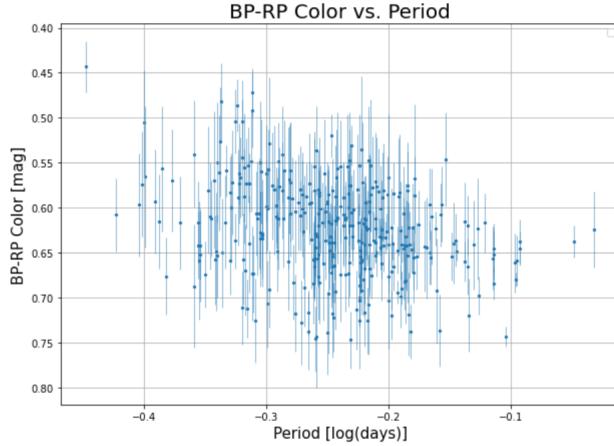


Figure 17: Color versus period plot of the data set of 500 clean RR Lyrae stars. This data will be inputted into the MCMC sampler for fitting.

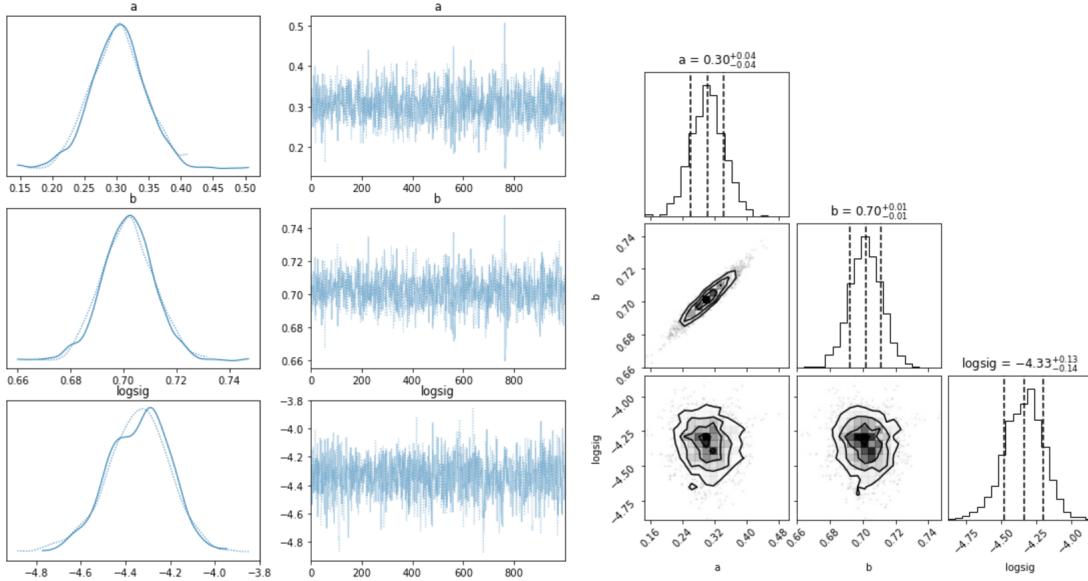


Figure 18: Trace and corner plots of the NUTS sampler fitting bp-rp color versus period.

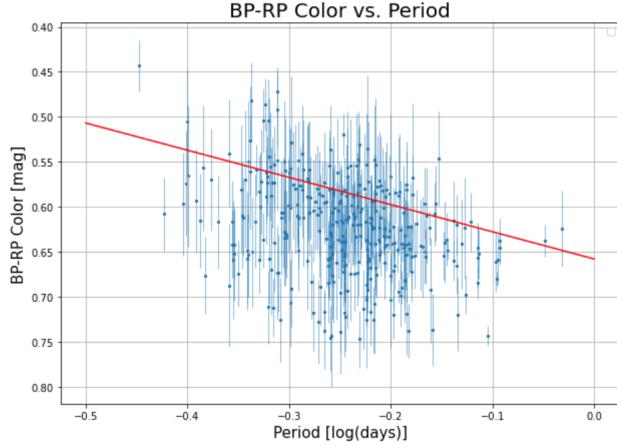


Figure 19: Best fit line using the mean of the posterior distributions of the parameters obtained from the NUTS sampler plotted against the same data as above. The parameters for this line are $a = 0.30$ and $b = 0.70$. The model is shifted upwards by a standard deviation to account for the fact that there shouldn't be any intrinsic scatter that causes sources to appear more blue. This model will be used to determine the intrinsic bp-rp color.

7 Constructing the galactic dust map

We now have a model for intrinsic bp-rp color as a function of period.

$$M_{bp-rp} = 0.30 * \log(P/\text{days}) + 0.70 \quad (13)$$

Given any RR Lyrae star, we can determine its intrinsic bp-rp color given its measured period. We now query every RRab star from the `vari_rrlyrae` catalog, regardless of its distance or galactic coordinates, cross matching it with the Gaia source catalog. We then calculate the extinction in the G band A_G , given by the equation:

$$A_G = 2.0 \times E(G_{BP} - G_{RP}) \quad (14)$$

where we are now able to solve for $E(G_{BP} - G_{RP})$ which is given in equation (1). We calculate A_G for each star and compare it with the `G_absorption` column in the `vari_rrlyrae` catalog.

As shown on Figure 18, there is a slight offset between `G_absorption` values and the calculated A_G values that may be due to further reddening that isn't accounted for in our initial derivation of intrinsic color. However, this is will just be an offset scaling factor and should not largely affect the structure of our constructed galactic dust map. To construct the dust map, we plot scatter plots of all the RR Lyrae stars in galactic coordinates against an Aitoff projection. We color each point by its normalized A_G value.

7.1 Comparison with the SFD map

The SFD map can be queried using the `dustmaps` package. To best compare our galactic map to the SFD map, we plot the same set of galactic coordinates and color them to the A_V values reported in the SFD query.

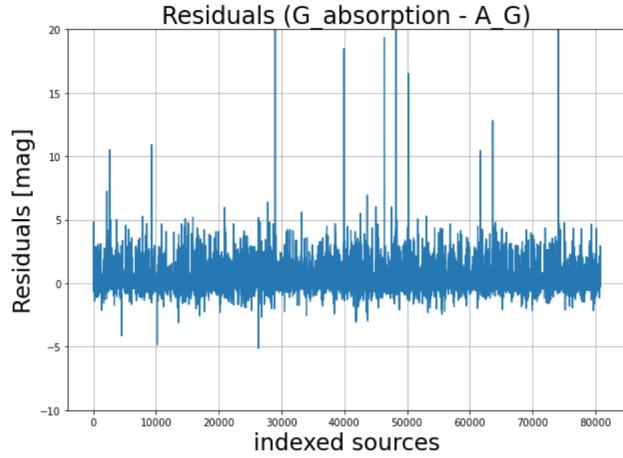


Figure 20: Residual plot between the reported `G_absorption` values and the calculated A_G derived from the color excess. There is a consistent deviation from the expected zero line. This could be due to the fact that the bp-rp values are not actually intrinsic and is actually affected by interstellar reddening, contrary to our initial assumptions.

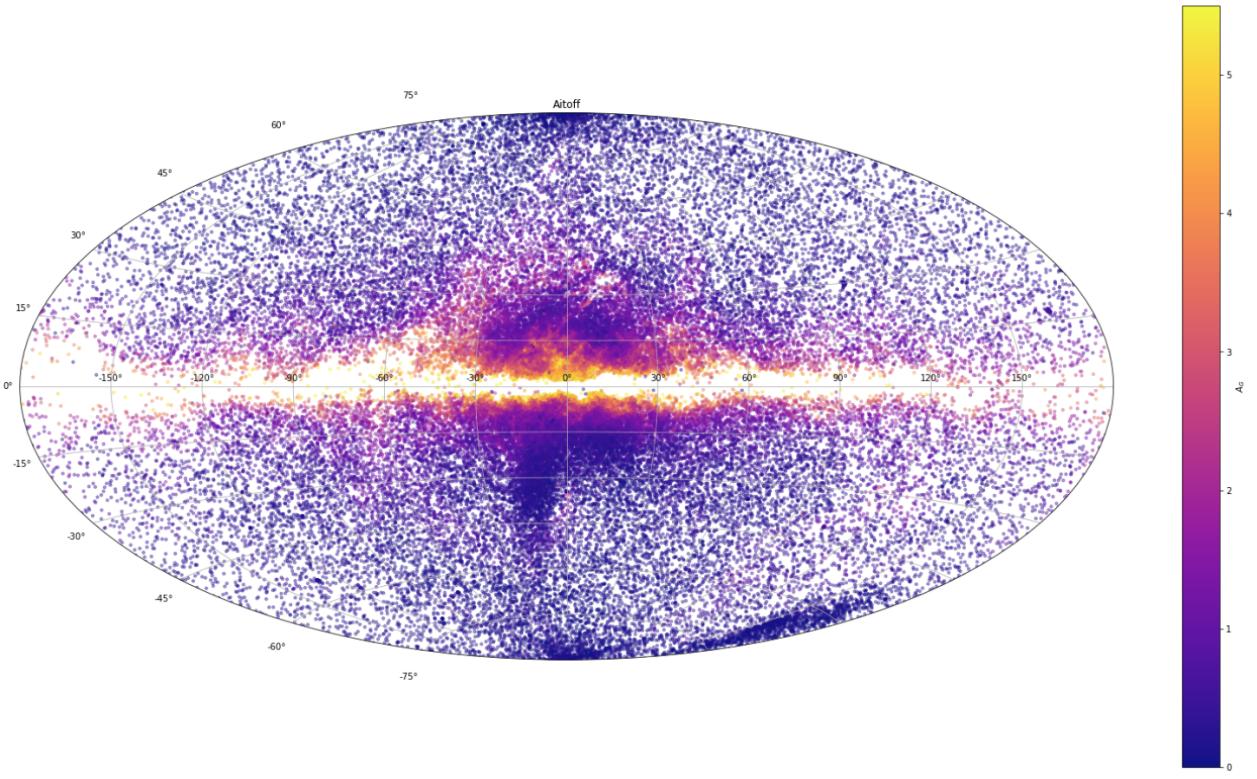


Figure 21: Constructed galactic dust map. Each point is a location of an RR Lyrae star colored by its A_G value. We can see that as the points get closer to the galactic disk, there is more extinction, as expected. The color range is arbitrary

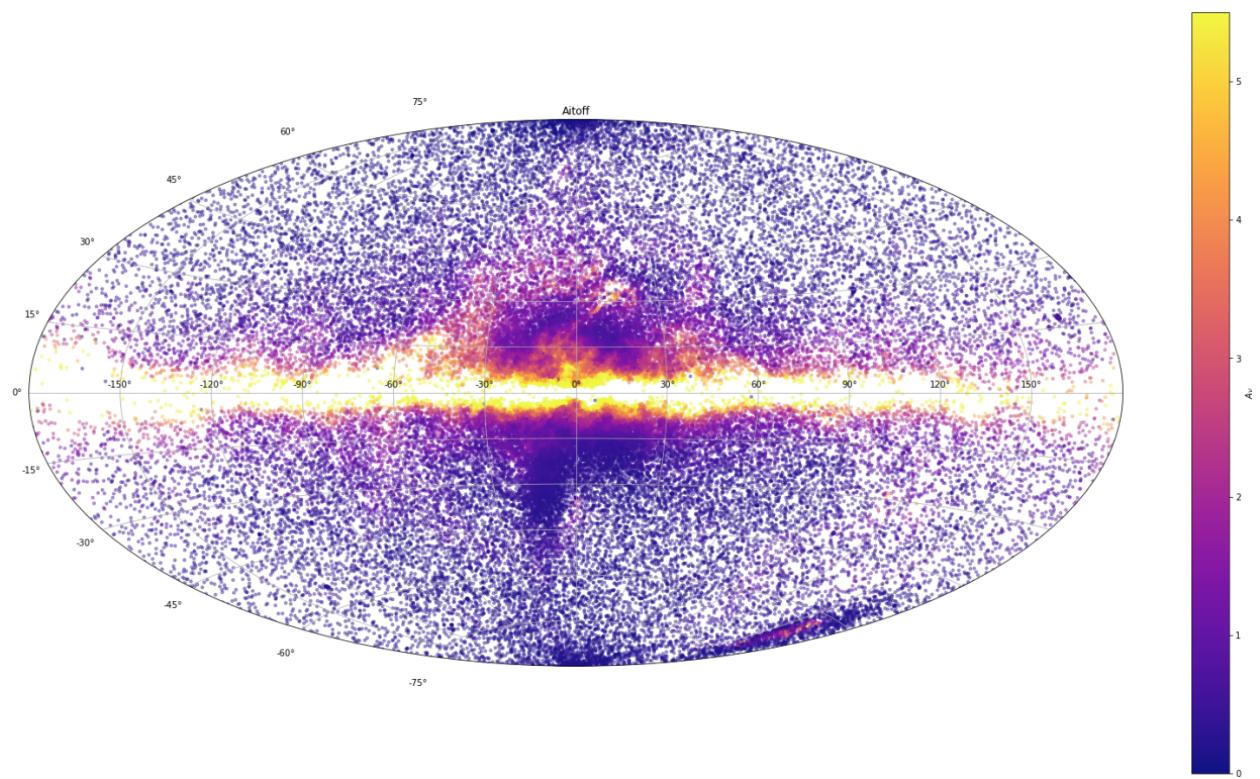


Figure 22: SFD map plotted at the same galactic coordinates of our constructed map. The general structure of the SFD map closely matches our constructed map. However, it appears to have less artefacts and thus looks smoother. There are much less points in which its extinction values do not match its surrounding values. The color range is arbitrary.

8 Conclusion

There is an overall matching general structure of the dust map, where extinction becomes more prominent near the galactic disk as well as the shapes near the edges of the disk. However, our constructed dust map contains many points in which its extinction values do not match its surrounding values. This is likely due to photometric noise or incorrectly measured periods, corresponding to a wrong intrinsic color. The SFD map was made by measuring the far infrared emission of dust. It is clear that a more direct observation of the dust will result in a more accurate and complete map as opposed to attempting to model a dust map in a passive manner by observing RR Lyrae stars.

A Appendix

References

- [1] C. A. L. Bailer-Jones et al. “Estimating Distances from Parallaxes. V. Geometric and Photogeometric Distances to 1.47 Billion Stars in Gaia Early Data Release 3”. In: *The Astronomical Journal* 161.3 (Feb. 2021), p. 147. DOI: 10.3847/1538-3881/abd806. URL: <https://doi.org/10.3847%2F1538-3881%2Fabd806>.
- [2] H Netzel et al. “Blazhko Effect in the first overtone RR Lyrae stars of the OGLE Galactic bulge collection”. In: *Monthly Notices of the Royal Astronomical Society* (July 2018). DOI: 10.1093/mnras/sty1883. URL: <https://doi.org/10.1093%2Fmnras%2Fsty1883>.
- [3] David J. Schlegel, Douglas P. Finkbeiner, and Marc Davis. “Maps of Dust Infrared Emission for Use in Estimation of Reddening and Cosmic Microwave Background Radiation Foregrounds”. In: *The Astrophysical Journal* 500.2 (June 1998), pp. 525–553. DOI: 10.1086/305772. URL: <https://doi.org/10.1086%2F305772>.

This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.