Finding the Period-Luminosity Relation of Cepheid Variables

Mark Ryan
Submitted: 18/04/2023

This paper presents an investigation into the period-luminosity relation of Cepheid variables by measuring the apparent V-band magnitude of 21 different Cepheids over an eight-week observation period. A two-harmonic function is then optimised to the data for each star, allowing the optimal period and average apparent magnitude to be obtained from this fit. Chi-square, χ^2, statistics are then used to assess the eligibility of the data for each of the stars and it is found that one star, CO Aur, has zero degrees of freedom and so must be discarded. The data for the remaining stars is then used in conjunction with parallax distance data from GAIA to calculate the absolute magnitude of each Cepheid and plot this against the period (which was obtained from the two-harmonic fits) to produce a period-luminosity graph. A linear model is then fitted to the data and fit parameters are obtained to find a period-luminosity relation of the Cepheids as $M = (0.6 \pm 0.1) - (2.8 \pm 0.1)\log(P)$, where M is the absolute magnitude and P is the period in days. According to χ^2 statistics, this model should be rejected as it yields a P-value less than 10^{-4} at 7.3×10^{-122} (2SF), alongside a reduced χ^2 value which was far from the desired value of one at 35 (2SF). Upon investigation into the cause of these errors, one finds that errors in the parallax distances may have been underestimated with systematic errors in these values not being accounted for. Furthermore, confidence in the two-harmonic fits for each star is low due to the small numbers of degrees of freedom caused by low numbers of observations, thus highlighting another potential source of error responsible for the poor χ^2 statistics which cause the model to be rejected.

1. Introduction

Cepheids are variable stars which pulsate radially and vary their brightness through changes in both their temperature and diameter [1]. These variations are periodic meaning that the pulsations yield stable periods and amplitudes which can be observed through monitoring the stars apparent magnitude. The light curves produced by monitoring these variations are sinusoidal and can be constructed through the summation of a number of harmonics, with more harmonics typically meaning that more small-scale variations to the periodic change in apparent magnitude can be accounted for. Through the use of a summation of the form,

$$m = m_0 + \sum_i a_i \sin\left(\left(\frac{2\pi n_i}{T_0}\right)t + \phi_i\right), \qquad (1)$$

the apparent magnitude, m, can be found at each observation time (in Modified Julian Days), t, by making use of the period of the Cepheid, T_0, and the average apparent magnitude, m_0. Using equation (1), it is easy to see how any number of harmonics can be implemented into fitting to the observational data, with each harmonic having a different amplitude, a, phase difference, ϕ, and multiple of $\left(\frac{2\pi}{T_0}\right)t$, n_i. Due to the nature of Cepheids, n_i, must be an even number for each harmonic, i, in order to ensure that the period of the oscillation remains as T_0.

Cepheids can have periods as high as 100 days, with Galactic Cepheids typically having a maximum of around 50 days [2]. Moreover, they can be up to hundreds of thousands of times the luminosity of the Sun meaning that in many cases, they can be observed at very large distances from Earth.

Cepheid variables follow a period-luminosity relationship by which as their period increases, so too does their luminosity. This characteristic was discovered by Henrietta Swan Leavitt in 1908 [3] and the relation is well defined meaning all Cepheids should follow the same empirical relation [4]. By taking a logarithmic scale for the period, one converts this proportionality into a linear relationship between the two variables [5],

$$M = \alpha \log_{10}(P) - \beta\ , \qquad (2)$$

where M is the absolute magnitude (which is analogous to luminosity with both of them representing intrinsic brightness), P is the period and α and β are constants which this investigation aims to find.

Once the values of α and β have been determined, Cepheids are able to be used as standard candles [6]. This is possible because if one measures the period of the apparent magnitude variation of a Cepheid, equation (2) can be used to find the value of its absolute magnitude. The actual value of absolute magnitude will need to be adjusted dependent on the wavelength filter desired, however this correction is just an offset captured in β, with the proportionality constant α being consistent across all filters.

After determining the absolute magnitude of the Cepheid, one can make use of the distance modulus equation,

$$m - M = 5\log_{10}(d) - 5\ \ , \qquad (3)$$

to find the distance to the Cepheid, d, in parsecs using the observed apparent magnitude averaged over a pulsation period, m, and the absolute magnitude, M, inferred from the period-luminosity relation. With this, one can move up the cosmic distance ladder and calibrate the next step of finding distances using Type 1a supernovae. These are extremely bright and have known magnitudes meaning that once the Cepheid period-luminosity relation has been used to calibrate the measurements of closer supernovae, more distant Type 1a supernovae can be used as standard candles to calculate the distances to further stellar objects [7]. Ultimately, this process can continue to allow progression up the cosmic distance ladder until Hubble's Law is determined. Once this point is reached, a value for the Hubble constant, H_0, will have been determined meaning the age of the universe can be inferred [8]. Currently, there is controversy surrounding the reliability of this method to find H_0 as using it produces slightly different values to determining H_0 using the cosmic microwave background radiation [9]. Therefore, by undertaking this investigation, one will be able to determine whether the claims that the method which makes use of standard candles is not as reliable as previously thought are true [10].

In this report, observations of the V-band apparent magnitude of 21 Cepheid variables are made over an eight-week period and plotted against the date and time of the observation (in Modified Julian Days). From here, two-harmonic fits are optimised to the data in order to find the period and an average apparent magnitude of each star. With these parameters, the absolute magnitude of each Cepheid is then calculated using parallax distances from GAIA [11] and plotted against the logarithm of the respective optimised period. A linear relation is then fitted to this data in order to calculate the parameters in the period-luminosity relation. Once the relationship is determined, chi-square, χ^2, analysis is performed to test the quality of the data before it is compared to period-luminosity relationships from literature. To close, conclusions are drawn about the data and experimental procedure.

2. Method

To begin the investigation, a list of 21 suitable Cepheid variable stars was constructed. In an attempt to eliminate any bias in the data and increase reliability of the results, the Cepheids were chosen such that there was a variety of predicted distances, magnitudes and pulsation periods across a suitable range.

With the ultimate goal of this investigation being to deduce a period-luminosity relationship of Cepheid variables, using Cepheids which cover the largest possible range of the aforementioned variables means that the extrapolation of this relation is more reliable and representative of the total population of Cepheids. This range was dependent on several factors. Firstly, the available stars had to be chosen based on if they were high enough above the horizon to be seen from the Durham observing site at a suitable time (approximately 7:00pm) for the entirety of the eight-week observation period – a high declination was prioritised here in order to reduce the atmospheric seeing in images taken [12]. Next, the maximum pulsation period was restricted by the eight-week time frame available for the investigation as it was imperative that at least one full pulsation period occurred within the time between the first and last observation of each star. Finally, the maximum and minimum magnitudes were dependent on the telescopes available – stars had to be of at least a minimum apparent magnitude such that they could be seen but less than some maximum magnitude such that they did not saturate the CCDs or prevent other stars being seen in the image. To further refine this list, the distance to each star was checked using data from GAIA [11] to ensure there was a reasonable variety.

The completion of this list saw 21 Cepheids with a predicted range of their periods of 1.78 – 26.1 days [13], a predicted range of apparent V-band magnitudes of 7.691 – 14.101 [13], and a predicted range of parallax distances of (0.95 ± 0.02) – (10 ± 2) kpc [11].

Table 1: A list of the 21 Cepheid variables chosen for this investigation alongside their predicted periods, apparent magnitudes and parallax distances. The periods and apparent magnitudes were obtained from the University of Toronto database [13] and the parallax distances have been obtained from the GAIA database [11].

Star	**Period (days)**	**Apparent Magnitude**	**Parallax Distance (kpc)**
AP Cas	6.847000	11.550	3.8 ± 0.2
RY Cas	12.138880	9.960	3.0 ± 0.1
CT Cas	3.810610	12.276	3.3 ± 0.2
SZ Cas	13.637747	9.862	2.7 ± 0.1
DD Cas	9.812027	9.894	3.1 ± 0.1
DL Cas	8.000669	8.984	1.81 ± 0.09
FW Cas	6.237050	12.449	4.3 ± 0.2
OT Per	26.091000	13.527	4.5 ± 0.5
BM Per	22.951900	10.461	3.5 ± 0.3
YZ Aur	18.193212	10.361	5.2 ± 0.4
RW Cam	16.414812	8.720	1.4 ± 0.3
AS Per	4.972516	9.760	1.63 ± 0.04
SY Aur	10.144698	9.095	2.3 ± 0.1
CO Aur	1.783010	7.715	0.95 ± 0.02
GP Per	2.041600	14.101	8 ± 1
EW Aur	2.659560	13.534	8 ± 1
CY Aur	13.847650	11.891	4.3 ± 0.4
RX Aur	11.623537	7.691	1.62 ± 0.06
ER Aur	15.690730	11.534	10 ± 2
HQ Per	8.637930	11.607	4.9 ± 0.6
BK Aur	8.002432	9.451	2.6 ± 0.1

Observations of the apparent V-band magnitudes of each of these stars then had to be taken as frequently as possible throughout the eight-week observation period. Before taking any images, a method that was used to achieve this was to focus the telescope, in the correct colour band, on a bright star until a seeing of approximately 3 arcseconds or less was achieved. Through doing this, any images taken on the telescope were improved as minimising seeing meant that as much of the turbulence from the atmosphere was eliminated. In order to improve said observations, one aimed to increase the signal-to-noise ratio (S/N) of the images taken. A method used to achieve this aim was to take more than one image for each observation and then stack the images to average out, and therefore reduce, the noise. Stacking images should, in theory, increase the S/N by a factor of $\sqrt{N}$, where N is the number of images being stacked [14].

Once the images were stacked, the zero-point magnitude was calculated by taking the median of all the zero-point magnitudes calculated using aperture photometry for each star in the image. The zero-point magnitude of an image is the magnitude of an object that produces one count per second on the detector [15]. The zero-point magnitude relative to each star in the image was calculated via,

$$m_{zpt} = m_1 + 2.5\log_{10}\left(\frac{f_1}{f_{zpt}}\right), \quad (4)$$

where the literature value of the apparent V-band magnitude of a known star in the image, m_1, was used alongside the count number (flux) corresponding to this star, f_1, to calculate the zero-point magnitude, m_{zpt}, which would correspond to a count rate of one count per second, $f_{zpt} = t_{exp}$. Here, t_{exp} is the exposure time of the image and it is possible to think of the flux in this way due to the flux being the product of the gain multiplied by the count rate (and since both are from the same image, the gain will be the same and it will cancel in the above equation – otherwise gain would be assumed to be one). The median of these values for each stacked image was taken in order to prevent any extreme values (from very bright stars for example) from dominating the average and the standard error was calculated as the error on this average.

To check that the expected increase in S/N was observed and decide how many images should be stacked for each observation, the error in the zero-point magnitude was plotted for different numbers of stacked images for observations of the star AP Cas.

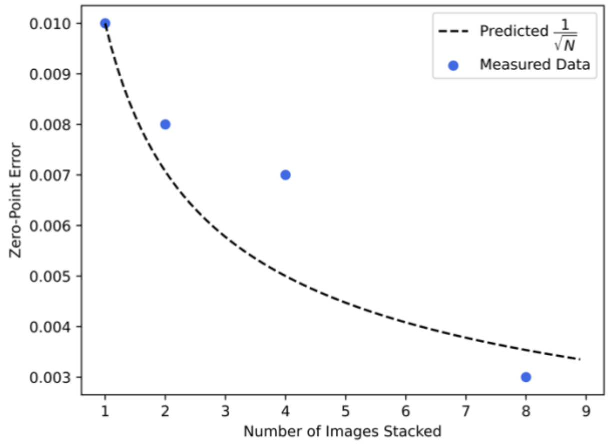


Figure 1: The error in the zero-point magnitude plotted against the number of images stacked. The expected trend is plotted in a blacked dashed line.

A clear decrease in the zero-point error as the number of images stacked increases can be seen in Figure 1, suggesting that stacking images does indeed improve the S/N, however, there is not a sufficient number of data points to be able to deduce whether the measured data follows the predicted trend. The trend in the zero-point error for different numbers of stacked images was able to be analysed to make conclusions surrounding the trend in the S/N because when stacking images, the signal will stay the same (since the stars are consistent throughout the images being stacked) but the noise will be averaged out, and thus decrease, since it is random across each image taken.

To calculate magnitudes from the images, including the aforementioned zero-point magnitudes, the optimal aperture size used to analyse the image had to be determined. Apertures were placed around the stars of interest so that the counts within them could be analysed to obtain the magnitudes of the stars based on their counts compared to the background count. To find this optimum size, the S/N and the total counts within the aperture had to be maximised and so both of these variables were plotted against different sized apertures.

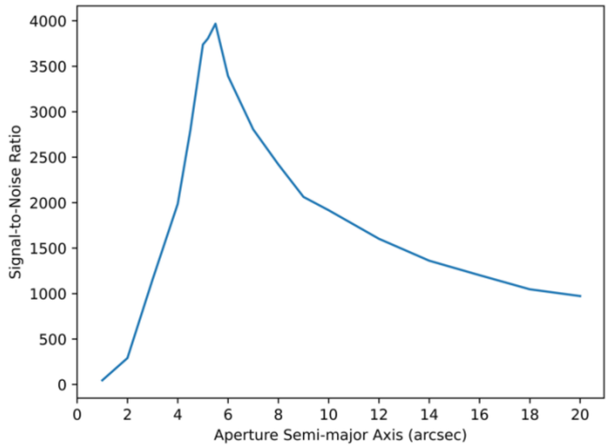


Figure 2: The S/N plotted against the semi-major axis of the aperture in arcseconds. The S/N peaks between 5-6 arcseconds suggesting a value in this range is optimal.

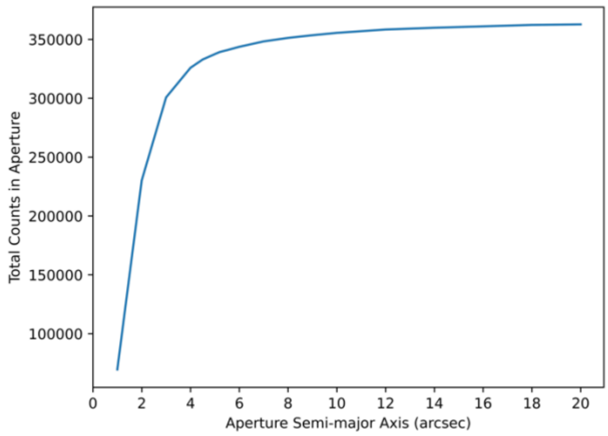


Figure 3: The total counts within the aperture plotted against the semi-major axis of the aperture in arcseconds. The number of total counts in the aperture starts to plateau after approximately 4-5 arcseconds suggesting that a value in this range is optimal as after this point the aperture is just capturing noise and so the total counts is only slightly increasing.

From Figures 2 and 3, the optimum aperture size for all magnitude analysis, including the zero-point magnitudes, was set to be 5 arcseconds as this is both where the S/N peaks and also where the total number of counts within the aperture begins to plateau.

Magnitudes were then extracted from each stacked image using apertures with a semi-major axis of 5 arcseconds. To calculate these magnitudes, an aperture of said size was drawn around the target star in the image. This aperture had an outer annulus which did not include any of the counts from the target star and within which the count rate was measured. The count rate between this outer annulus and the inner aperture was the background count.

Averaging the sky background count out per pixel then allowed for the count from the target star to be found by summing the total counts within the aperture and subtracting the product of the background count per pixel with the area of the aperture. This value was then converted to a magnitude using a rearranged version of equation (4),

$$m_{target} = m_{zpt} - 2.5\log_{10}\left(\frac{f_{target}}{f_{zpt}}\right), \tag{5}$$

where now the magnitude of the target star, m_{target}, could be calculated using the previously calculated zero-point magnitude, m_{zpt}, and the measured counts for the star and zero-point, f_{target} and f_{zpt}, respectively.

Each magnitude had an associated error which was found using photon statistics as this offered the highest error on the value. When calibrating the magnitudes with the zero-point, the aforementioned error on the zero-point had to be accounted for and so to deduce an error on the final magnitudes of each target star, the error on the apparent magnitude using photon statistics was added in quadrature with the error on the zero-point magnitude. In each case, the error on the zero-point magnitude was dominant.

From here the apparent V-band magnitudes were plotted against time for each of the 21 stars and a two-harmonic periodic function was fitted to the data. Normalised residuals were then able to be calculated and plotted for each fit. A two-harmonic fitting function was chosen over a different number of harmonics as when compared with single and three-harmonic fits, the chi-square, χ^2, statistics were optimised.

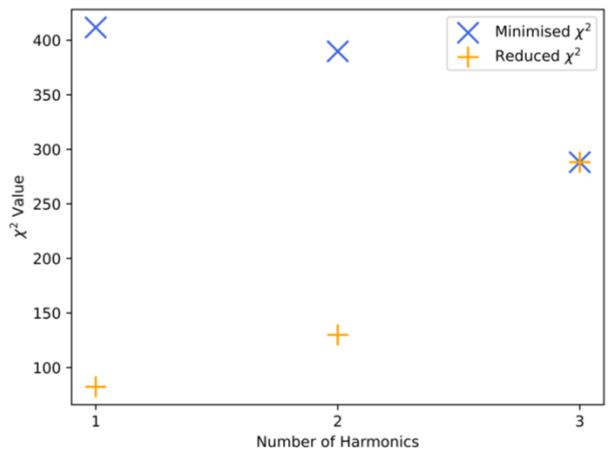


Figure 4: Values of minimised χ^2 and reduced χ^2, for one of the stars. The second harmonic was deemed to be optimal as it yields the second smallest value in both the minimised and reduced cases while the one-harmonic case has the highest minimised χ^2 and lowest reduced χ^2 with the three-harmonic yielding the converse.

The two-harmonic fits enabled optimal parameters for the period of the star and the average magnitude to be found after supplying an initial guess. Initial guesses of the period of each star were decided using a binning function on data from ASAS-SN [16], with any stars whose data was not sufficient for the binning function to produce an appropriate value having their initial guess taken as their predicted period from literature [13].

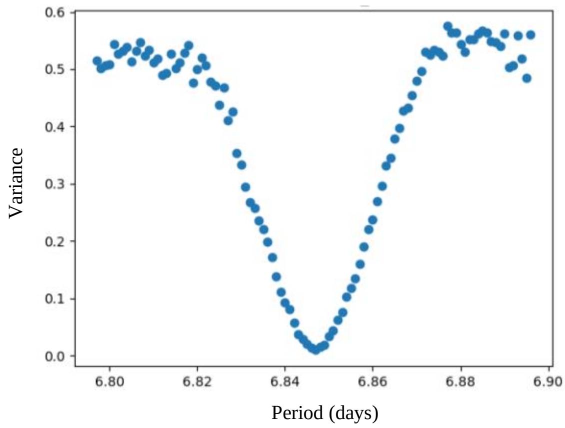


Figure 5: An example of the binning function used on each star to find an initial guess for the period. This figure shows the function being used on AP Cas and finds its optimal period to be 6.847 days.

After the initial guesses for each star were determined, the two-harmonic fits were plotted for each Cepheid and using the optimised parameters obtained from these, the period-luminosity relationship was able to be calculated. With luminosity being analogous to absolute magnitude, the absolute magnitude, M, of each star was calculated, using the median apparent magnitude, m_{av}, (calculated based on the minimum and maximum magnitude of the two-harmonic fit) alongside the parallax distance, d, in parsecs from GAIA [11] via,

$$M = m_{av} - 5\log_{10}(d) + 5\,. \tag{6}$$

This then allowed the absolute magnitude to be plotted against the logarithm of the periods that were obtained from the optimised two-harmonic fits and gave the period-luminosity relationship of the Cepheids in question.

Errors on both the absolute magnitude and the logarithm of the period then had to be calculated so that they could be added to the plot. To calculate the errors in the absolute magnitude, α_M, the functional approach to error analysis [17] was taken in order to find the contributions to α_M from the errors on each of the variables used to calculate the absolute magnitude ($\alpha_{M,m_{av}}$ and $\alpha_{M,d}$). These contributions were calculated using the general function,

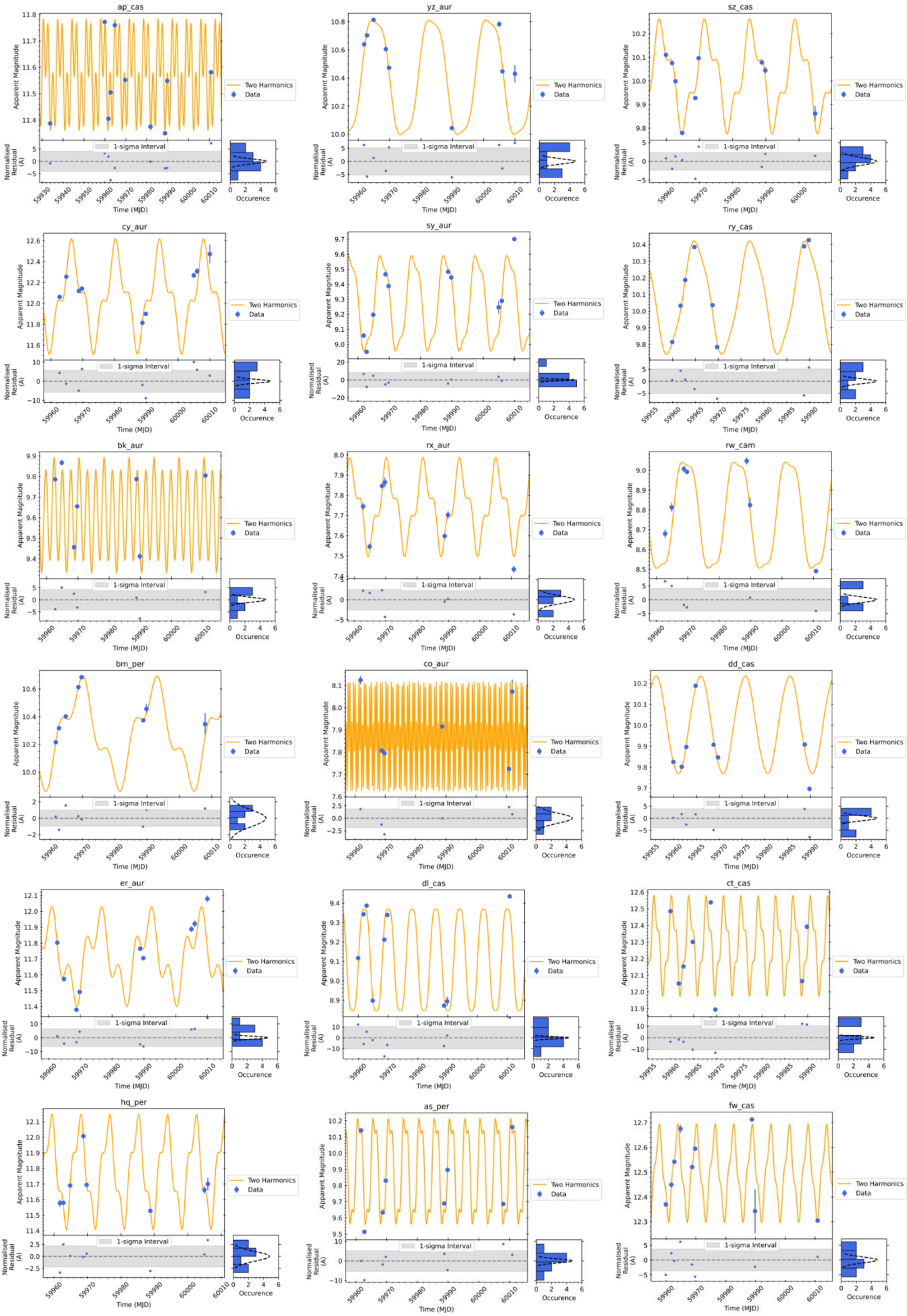
$$\alpha_{f,u} = |f(u+\alpha_u, w, \dots) - f(u, w, \dots)|\,, \tag{7}$$

where $\alpha_{f,u}$ represents the contribution from the error of a specific variable, u, in the function. This was used to calculate $\alpha_{M,m_{av}}$ and $\alpha_{M,d}$, before these contributions were added in quadrature to give the total error on the absolute magnitude values, α_M.

In accordance with the literature trend [5], a linear trendline was fitted to the data and normalised residuals were calculated and plotted to test the quality of the fit. Chi-square statistics were then used to further analyse the fit and test its quality, before the relation was further compared to literature.

3. Results & Discussion

Once data collection had been completed, each star had its own apparent V-band magnitude plotted against the time at which the observation was taken. This time was converted to Modified Julian Days (MJD) in order to give each observation a numerical value corresponding to the exact date and time at which it was taken. Doing this ensured that the results were as precise as possible as the precision (of the nearest second) of the timing instrument on the telescope was maintained.


ap_cas
yz_aur
sz_cas
cy_aur
sy_aur
ry_cas
bk_aur
rx_aur
rw_cam
bm_per
co_aur
dd_cas
er_aur
dl_cas
ct_cas
hq_per
as_per
fw_cas
Apparent Magnitude
Two Harmonics
Data
Normalised Residual (A)
1-sigma Interval
Time (MJD)
Occurence

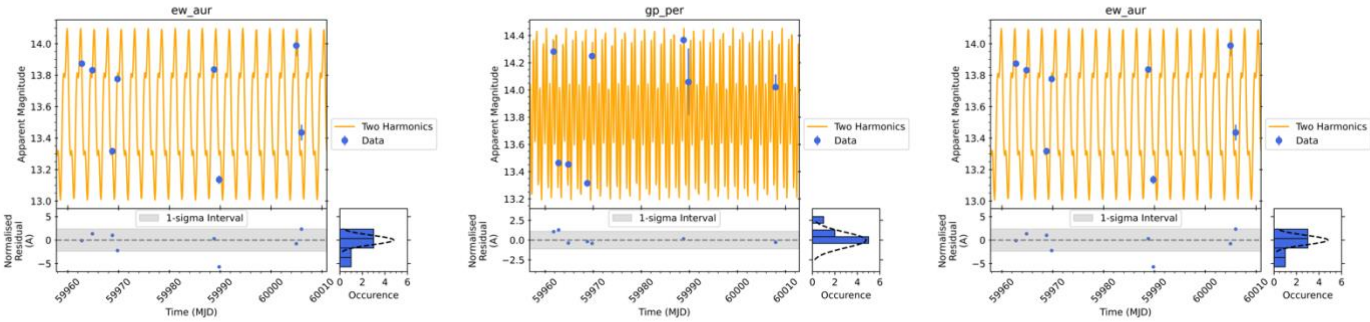


Figure 6: Observed apparent V-band magnitude at different times throughout the eight-week observation period for each of the 21 stars. The names of the stars can be found above each graph. Normalised residuals have been plotted with a shaded band showing the one standard deviation interval from the ideal mean of zero, as well as histograms representing each plot respectively. Both vertical and horizontal error bars have been plotted on each measured data point.

Using the two-harmonic fits in each of the graphs in Figure 6, optimal values of the period and average values of the apparent V-band magnitude of each of the Cepheids were deduced using a curve fitting software. Errors on these values were calculated using covariance matrices.

Each of the optimised fits were then tested for their quality using chi-square statistics. Using this meant that one could obtain an insight into how reliable the fits were to the data set that had been collected. One of the biggest issues with the fitting functions used is that much of the fit had to be extrapolated and estimated in between data points due to their being large gaps of up to two weeks between consecutive observations at times. Therefore, assessing the quality of each of the fits would aid in diagnosing any future problems that might become prevalent when the parameters obtained from said graphs were used to plot the period-luminosity relation.

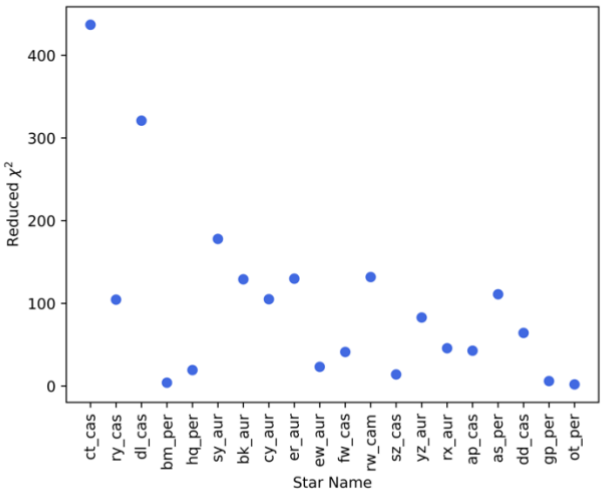


Figure 7: Reduced χ^2 plotted for each of the stars observed. The only star which does not appear on this plot is CO Aur due to it having zero degrees of freedom meaning that reduced χ^2 could not be calculated for said Cepheid. The mean value of reduced χ^2 for the 20 eligible stars was 99.7 (3SF) which is larger than the desired value of one suggesting that on average, the fits are not sufficient.

Despite analysis of χ^2 implying that many of the fits are poor overall, Figure 7 shows that it is suitable to use the data for each of the stars except for CO Aur since it has zero degrees of freedom. Having zero degrees of freedom means that there is no way to affirm or reject the model since the data has no freedom to vary and so one cannot conduct research with this data set. As a result of this, the data for CO Aur was discarded in order to avoid jeopardising the validity of the period-luminosity relation and therefore, only the data for the other 20 Cepheids was used. This decision was made as although some of the fits yield poor χ^2 statistics, they are still valid and can be used in research so discarding any other points would risk damaging the integrity of this investigation.

Using the optimised fit parameters for the 20 eligible stars, the absolute magnitude was plotted against the logarithm of the period to produce a period-luminosity relation of Cepheid variables.

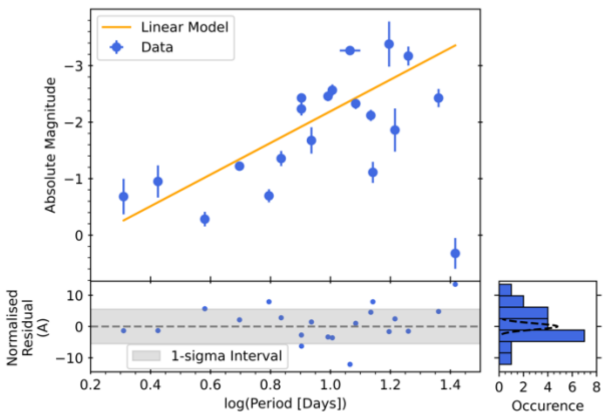


Figure 8: Absolute magnitude of each eligible star plotted against a log scale of the observed periods of each of the stars. A linear model relating the two variables has been plotted. Normalised residuals have been plotted with a shaded band showing the one standard deviation interval from the ideal mean of zero, as well as histograms representing each plot respectively. Both vertical and horizontal error bars have been plotted on each measured data point.

Using the errors associated with the absolute magnitudes of each star, a linear model was plotted to fit the trend of the data. A linear relationship was expected based on literature [5] as this type of relationship allows extrapolation which is useful when using the period-luminosity relationship of Cepheids to move up the cosmic distance ladder [8]. Optimising the linear fit allowed fit parameters and their associated errors to be calculated (again using a method that made use of covariance matrices). This investigation yielded a period-luminosity relation of $M = (0.6 \pm 0.1) - (2.8 \pm 0.1)\log(P)$ which insinuates that as the period of Cepheids increases, the intrinsic brightness of the stars also increases (thus the absolute magnitude of the stars decreases as observed in Figure 8).

Performing χ^2 analysis on this data set with the optimised linear fit found that the model was a poor fit to the data yielding a reduced χ^2 value of 35 (2SF) which is far from the ideal value of one (despite its 18 degrees of freedom), which suggests that this model should be questioned, even though it follows the expected trend. Further analysis of this model reveals that it should in fact be rejected as its P-value is much less than 10^{-4} [17] at 7.3 x 10^{-122} (2SF). This is rather surprising considering a linear relationship was expected and so it implies a potential error within the observations or grossly underestimated errors on said observations.

To investigate the problems within this data the normalised residuals were investigated in the hope that any problems surrounding the errors could be identified and diagnosed. The mean of the normalised residuals was 1.0 (2SF) which is far from the ideal mean of zero. Moreover, the residuals seem to have a somewhat heteroscedastic nature with no other obvious trend being visible in them. The spread of these points however is fairly large with the residuals having a standard deviation of 5.5 (2SF) and they do not quite follow the desired normal distribution as seen in the histogram in Figure 8. The overall pattern seen in this histogram is relatively close to a normal distribution suggesting that the problem is likely due to underestimates in the error which have caused the spread in the residuals to be amplified. This is quite a positive observation as it suggests that the problem lies with the error estimates rather than the observations themselves and thus hints at potential reliability in the trend of data set. This notion also fits with the poor χ^2 statistics, further suggesting that there is an issue with the error values for the absolute magnitudes.

A problem with the absolute magnitude calculation is also supported by the anomalous result produced by OT Per. This point can be seen in the bottom right of the period-luminosity plot found in Figure 8. Since this point lies over 12 standard errors from the optimised linear fit line, there is a strong suggestion that there is either an error in the optimised period (and its error) used or in the absolute magnitude (and its error) calculated.

To investigate the issue in the absolute magnitude error, the components to calculate this error needed to be considered separately. The absolute magnitude had contributions from the average apparent magnitude and the parallax distance to the star. Since apparent magnitude is dependent on the seeing that the observation is subject to, taking a literature value for this for each star would be difficult. Therefore, since the method of the average apparent magnitudes' calculation (alongside its error) was identical to that used to calculate the period and its error, the accuracy of the optimised period was analysed. The periods of the Cepheids will be relatively consistent regardless of where the observations were taken and what equipment was used meaning that analysis of the calculated period against the known period [13] should give an insight into the reliability of the optimisation method used to calculate this variable and its error. In doing this, one also learns about the reliability of the average apparent magnitude and its error obtained from the same optimisation method, thus aiding in deciding the dominating factor in the poor absolute magnitude error calculations.

For this investigation, an O-C diagram was plotted such that the "calculated" (literature) period was taken away from the "observed" (optimised) period and plotted for each star to see its deviation from zero.

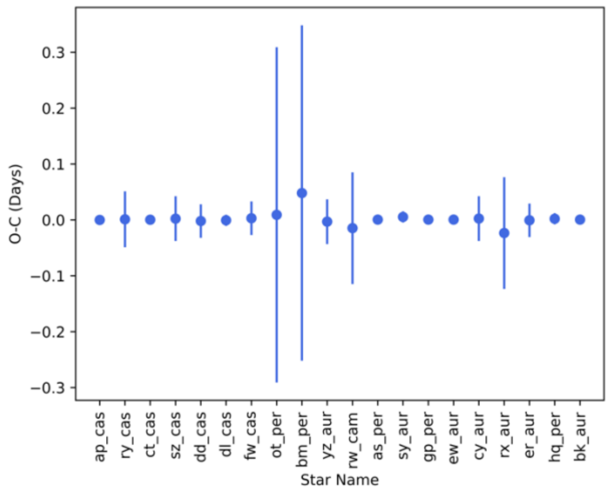


Figure 9: O-C diagram representing the difference between the period obtained from the optimised two-harmonic fits and the period found in literature for each of the 20 eligible Cepheids. Each point has vertical error bars which are the errors generated using the covariance matrices when optimising the fit parameters.

As seen in Figure 9, all of the observed periods which were found by optimising the two-harmonic fit to each star's dataset were very close to the known periods found in literature [13], with the mean difference between these two values being 0.0015 days (2SF). The trend seems constant which suggests that there are no systematic errors in the calculation of the period. For each of the stars, the known period lies within one standard error of the observed period meaning that the optimisation method used to find the periods produces results in strong agreement with literature in 100% of cases (as seen in Figure 9 where all error bars include zero as the difference between the two values). Therefore, one could argue that this implies that the error calculation for the periods is suitable as the error range is large enough to cause an excellent agreement, which further insinuates that the method used to calculate these errors is suitable. Since the same method was used for both the period and its error and the average apparent magnitude and its error, this analysis suggests that the average apparent magnitude calculation and its error should also be suitable and trusted. As a result of this, one can deduce that the problem with the absolute magnitudes and their errors should not be stemming from the contribution to the error from the average apparent magnitude (and its error) calculations.

With the absolute magnitude being calculated using both the average apparent magnitude and the parallax distance, the above deductions suggest that the problems with the errors in the absolute magnitudes lie within the parallax distances and their errors. The parallax distances and their errors were obtained from the GAIA archive provided by the European Space Agency (ESA) [11]. Researching the quality of this data revealed that the reliability of the errors quoted on the parallax values is very much in question [18]. The precision in the GAIA parallax data has been questioned for a long time with even earlier data releases being questioned for its errors [19], causing the ESA to have a webpage discussing potential problems with the parallax data on which they suggest some unaccounted for systematic and random errors may be present in the data [20]. Therefore, if there are additional errors in the parallax data which have not been included, then the actual error in the values may be much larger than that which has been used in the calculations for the absolute magnitudes of each of the stars. These problems align well with the issues found in the data in this investigation supporting the claim that the poor χ^2 statistics for the period-luminosity relation are due to underestimated errors for the absolute magnitude which stem from the contribution to this error from the parallax distances. To attempt to improve the data, one would have to find a method of quantifying the systematic errors in the parallax distances. One method of doing this would be to map the systematic errors using analysis of a large number of Milky Way stars [21], however it would be a huge task to investigate a sufficient number of stars which gave confidence in there being any systematic uncertainties hidden in the parallax data.

Furthermore, other methods to calculate distances to Cepheids could also be tested to examine whether they yielded reproducible results. One potential method to do this would be to use a near-infrared variant of the Barnes-Evans method for finding Cepheid distances [22]. This could then be compared to results found using GAIA's parallax data to infer whether there is in fact a problem with the distance data used.

Despite the errors in the data discussed, when comparing the relation to literature relations for Cepheids one finds a variety of results with some papers like that of Molinaro et al. in 2011 [23] and Turner in 2010 [24] yielding relations which have slopes (-2.78 ± 0.11 and -2.78 ± 0.12 respectively) within one standard error of the one from the investigation discussed in this paper implying a strong agreement, and other papers like Benedict et al. in 2007 [25] and Storm et al. in 2011 [26] whose gradients (-2.43 ± 0.12 and -2.67 ± 0.10) are in disagreement with the one discussed in this paper, lying within four and two standard errors of it respectively. The fact that the data is in agreement with some literature, and it is not too far from the literature values that it is not in agreement with is positive and further supports the idea that the main issue with this data is the underestimation of errors, rather than the data itself.

The intercepts of each of these relations are in disagreement with each other however it was the slope of the period-luminosity relation which was focussed on as opposed to the intercept because the intercept of the relation is just an offset which can be due to the type of filter used, the seeing at the observation site and even systematic errors due to the equipment and calibration methods used. This means that the difference in offset does not need too much focus when initially comparing the quality of the relation as for this investigation, the proportionality of the absolute magnitude and period is more important.

Having a reliable and accurate value for the slope of the period-luminosity relation of Cepheids has many wider implications in the field of astronomy, such as finding distances to galaxies for example. Combining this with other experimental methods, such as finding recessional velocities using redshifts of galaxies, would ultimately allow for a value for the Hubble constant, H_0, to be determined [27]. Recently there has been widespread controversy over the value of this constant due to different methods of determining it yielding slightly different values. The discrepancy in the gradient obtained in this investigation with those from literature mentioned above adds to this controversy as the reproducibility of results from an investigation into Cepheids seems relatively poor. With the literature values obtained from a much larger investigation than the one detailed in this report not yielding one agreed value for the constant of proportionality within the period-luminosity relation of Cepheids, the reliability of this method comes into question. This idea is supported by literature claims that the period-luminosity relation is not as reliable as previously thought [28] and so could be reasoned to be the explanation as to why using standard candles to find H_0 [8], and thus the age of the universe, does not produce the same results as if one used the cosmic microwave background radiation to find said values.

As a result of this, further investigation of more Cepheids and more consistent observations over a longer observation period would be useful to be able to deduce whether or not the discrepancy in the period-luminosity relation's reliability is caused by GAIA's parallax data [18]. If multiple independent investigations which did not neglect systematic errors from GAIA were conducted and did not produce relations which had slopes that were all in agreement, there would be a strong claim to suggest that this method of calculating distances on the cosmic distance ladder is not as reliable as previously thought.

5. Conclusions

To conclude, a suitable list of 21 Cepheid variables which had a range of periods, magnitudes and parallax distances was constructed. These were then successfully observed over the eight-week observation period and a two-harmonic fit was optimised to each star's data. This optimisation process required an initial guess for the period which was determined in most cases using a binning function to find the optimal period based on ASAS-SN data [16]. χ^2 analysis on these fits revealed that the data of one star, CO Aur, had to be discarded due to it having zero degrees of freedom. The two-harmonic fits returned optimised values for the period and average apparent magnitudes of each star, and these were used to produce a period-luminosity relation for the Cepheids. The relation obtained was $M = (0.6 \pm 0.1) - (2.8 \pm 0.1)\log(P)$.

Despite showing the expected proportionality, the calculated period-luminosity relation had a poor reduced χ^2 value which was far from the desired value of unity at 35 (2SF). According to χ^2 statistics, the relation obtained should be rejected, yielding a P-value much less than 10^{-4} [17] at 7.3 x 10^{-122} (2SF). This was surprising due to the correct correlation being found, thus suggesting that the errors within this investigation must have been grossly underestimated.

Investigation into this issue found no systematic errors in the period, with all of the known periods from literature showing strong agreement with the optimised period and lying within one standard error. This implied that the optimisation had no systematic errors and so suggested that the optimisation method to calculate the period and its error was suitable. Since this method was also used to calculate the optimal average apparent magnitude for each star and its error, it was inferred that these values should too be accurate. This left the only obvious source of error as the parallax distances which were obtained from GAIA [11]. Upon investigation, the likelihood that the problem lay within the parallax distance measurements grew, with many sources stating undiagnosed systematic errors were hidden within the parallax distance data [18].

The relation was then compared to several different relationships from literature and found that it was in strong agreement with two of the chosen sources, Molinaro et al. in 2011 [23] and Turner in 2010 [24], despite the poor χ^2 statistics. Further comparisons found that it was also in disagreement with some literature relations like Benedict et al. in 2007 [25] and Storm et al. in 2011 [26], for example.

With neither χ^2 analysis nor comparisons to literature yielding convincing arguments to suggest that the relation obtained from this investigation is correct, one must conclude that the controversy in the reliability of the method of calculating H_0 using the period-luminosity relationship of Cepheids and standard candles [10] is valid and other methods like using the cosmic microwave background radiation may be more reliable [9]. Despite having a gradient that agrees with some literature, the poor χ^2 statistics suggesting that the model isn't accepted means that trusting this relation would show a great negligence to the uncertainties of the data which would not be acceptable.

To improve this investigation, one would look into using a method to calculate to calculate the systematic errors in the parallax uncertainties [21] in order to improve the absolute magnitude calculations and thus improve the reliability of the period-luminosity relation. Moreover, repeating the experiment over a longer time period with more consistent observations would allow the two-harmonic fits to be more accurate as the optimisation process would have more points to base the fit on. Due to the low number of points, many of these fits had a low number of degrees of freedom, with one of the stars (CO Aur) having zero causing it to be discarded. Therefore, by increasing the number of observations, the number of degrees of freedom in the data would be increased thus improving the fit quality and parameters obtained from these fits. It would also allow a higher number of harmonics to be plotted as the increased number of degrees of freedom would mean it was possible to have more free parameters. Through doing this, one could again compare the obtained relations to literature and investigate the χ^2 statistics to reassess the reliability of the method to calculate H_0 which makes use of standard candles.

References

[1] B. F. Lane et al, "Long-Baseline Interferometric Observations of Cepheids", *The Astrophysical Journal*, 573:330-337, July 2002

[2] W. P. Gieren et al, "Calibrating the Cepheid Period-Radius Relation with Galactic and Magellanic Cloud Cepheids", *The Astrophysical Journal*, 512:553-557, February 1999

[3] P. Murdin, "Leavitt, Henrietta Swan (1868-1921)", *Encyclopedia of Astronomy and Astrophysics*, 1st Edition, CRC Press, 2000

[4] L. Breuval, "The Cepheid Distance Scale: from the Local Gaia Calibration to Distant Galaxies", *PhD Thesis, University PSL*, October 2021

[5] J.D. Fernie, "A New Approach to the Cepheid Period-Luminosity Law: (Delta) Scuti Stars as Small Cepheids", *Astronomical Journal*, vol. 103, p.1647, May 1992

[6] N. R. Tanvir, "Cepheid Standard Candles", *Harmonizing Cosmic Distance Scales in a Post-Hipparcos Era, ASP Conference Series*, vol. 167, p. 84-100, 1999

[7] D. Branch, G. A. Tammann, "Type Ia Supernovae as Standard Candles", *Annual Review of Astronomy and Astrophysics*, vol. 30:359-389, September 1992

[8] B.R. Zhang et al, "A blinded determination of H_0 from low-redshift Type Ia supernovae, calibrated by Cepheid variables", *Monthly Notices of the Royal Astronomical Society*, vol. 471, p.2254-2285, October 2017

[9] J. L. Bernal et al, "The trouble with H_0", *Journal of Cosmology and Astroparticle Physics*, Issue 10, October 2016

[10] A.G. Riess et al, "Milky Way Cepheid Standards for Measuring Cosmic Distances and Applications to *Gaia* DR2: Implications for the Hubble Constant", *The Astrophysical Journal*, 2018

[11] Gaia Collaboration et al, (2016b): The Gaia mission (provides a description of the Gaia mission including spacecraft, instruments, survey and measurement principles, and operations); (2022k): Gaia DR3: Summary of the contents and survey properties

[12] H. T. Intema, “The GMRT 150 Mhz all-sky radio survey”, *Astronomy and Astrophysics*, vol. 598, A78, February 2017
[13] J. D. Fernie, B. Beattie, N. R. Evans, S. Seager, “Database of Classical Cepheids”, *International Bulletin of Variable Stars*, No. 4148, 1995 (Accessed: January 2023)
[14] J. Xu et al, “Minimum semblance weighted stacking with polarity correction for surface microseismic data processing”, *The Leading Edge*, 38 (8):630-636, August 2019
[15] C. N. A. Willmer, “The Absolute Magnitude of the Sun in Several Filters”, *The Astrophysical Journal Supplement Series*, vol. 236, No. 2, June 2018
[16] Shappee et al. (2014); The ASAS-SN Catalog of Variable Stars I: Jayasinghe et al. (2018a) (Accessed: January 2023)
[17] I. G. Hughes and T. P. A. Hase, *Measurements and Their Uncertainties*, Oxford University Press: Oxford (2010)
[18] K.A. Owens, W.L. Freedman, B.F. Madore, A.J. Lee, “Current Challenges in Cepheid Distance Calibrations Using Gaia Early Data Release 3”, *The Astrophysical Journal*, vol. 927, p.18, March 2022
[19] X. Luri et al, “Gaia Data Release 2”, *Astronomy and Astrophysics*, vol. 616, A9, August 2018
[20] *ESA Science & Technology – Known Issues with the GAIA DR2 Data*, (Accessed: March 2023), Available at: https://www.cosmos.esa.int/web/gaia/dr2-known-issues#AstrometryConsiderations
[21] M. A. Fardal et al, “Mapping Gaia Parallax Systematic Errors over the Sky with Faint Milky Way Stars”, *The Astronomical Journal*, vol. 161:58 (14pp), February 2021
[22] C. D. Wilson et al, “The Distance to IC 10 from Near-Infrared Observations of Cepheids”, *The Astronomical Journal*, vol. 111, October 1996
[23] R. Molinaro et al, “CORS Baade-Wesselink Method in the Walraven Photometric System: The Period-Radius and the Period-Luminosity Relation of Classical Cepheids”, *Monthly Notices of the Royal Astronomical Society*, vol. 413, Issue 2, pp. 942-956, May 2011
[24] D. G. Turner, “The PL Calibration for Milky Way Cepheids and its Implications for the Distance Scale”, *Astrophysics and Space Science*, 326:219-131, January 2010
[25] G. F. Benedict et al, “Hubble Space Telescope Fine Guidance Sensor Parallaxes of Galactic Cepheid Variable Stars: Period-Luminosity Relations”, *The Astronomical Journal*, vol. 133, Issue 4, pp. 1810-1827, April 2007
[26] J. Storm et al, “Calibrating the Cepheid Period-Luminosity Relation from the Infrared Surface Brightness Technique”, *Astronomy and Astrophysics*, vol. 534, A94, October 2011
[27] M. Abdel-Sabour et al, “Determination of the Hubble Constant Using Cepheids”, *Romanian Astronomical Journal*, Vol. 19, p.35-50, 2009
[28] M.R. Becker et al, “On the Impact of Cepheid Outliers on the Distance Ladder”, 2015

Appendix I
Observation Log

Table 2: Observation log for the investigation. Star names, date of observation and the telescope used for each observation are clearly shown. If an observation was made for a particular star on a particular date, the telescope which took the observation is named in that box. Any boxes marked with "-" are days when no observations for that star were taken, or if they were taken, they were not used in the investigation.

	15/01/23	**17/01/23**	**18/01/23**	**20/01/23**	**24/01/23**	**25/01/23**	**06/02/23**	**13/02/23**	**14/02/23**	**01/03/23**
AP Cas	Draco2	East-16	Draco2	West-14	-	Draco2	East-16	East-16	West-14	-
RY Cas	Draco2	East-16	Draco2	West-14	Draco2	Draco2	-	Draco2	West-14	-
CT Cas	Draco2	East-16	Draco2	West-14	Draco2	Draco2	-	East-16	West-14	-
SZ Cas	Draco2	East-16	Draco2	West-14	Draco2	Draco2	-	Draco2	West-14	Pt5m
DD Cas	Draco2	East-16	Draco2	West-14	Draco2	Draco2	-	Draco2	West-14	-
DL Cas	Draco2	East-16	Draco2	West-14	Draco2	Draco2	-	Draco2	West-14	-
FW Cas	Draco2	East-16	Draco2	West-14	Draco2	Draco2	-	East-16	West-14	-
OT Per	-	East-16	Draco2	West-14	Draco2	Draco2	-	East-16	Draco2	Pt5m
BM Per	-	East-16	Draco2	West-14	Draco2	Draco2	-	Draco2	Draco2	-
AS Per	-	East-16	Draco2	West-14	Draco2	Draco2	-	Draco2	Draco2	-
GP Per	-	East-16	Draco2	West-14	East-16	Draco2	-	-	Draco2	-
HQ Per	-	East-16	West-14	West-14	East-16	Draco2	-	Draco2	-	Pt5m
RW Cam	-	-	West-14	West-14	Draco2	Draco2	-	Draco2	Draco2	-
YZ Aur	-	East-16	West-14	West-14	Draco2	Draco2	-	-	Draco2	Pt5m
SY Aur	-	East-16	West-14	West-14	Draco2	Draco2	-	Draco2	Draco2	Pt5m
CO Aur	-	East-16	-	-	Draco2	Draco2	-	Draco2	-	-
EW Aur	-	-	West-14	West-14	East-16	Draco2	-	East-16	Draco2	Pt5m
CY Aur	-	-	West-14	West-14	East-16	Draco2	-	Draco2	Draco2	Pt5m
RX Aur	-	-	West-14	West-14	Draco2	Draco2	-	Draco2	Draco2	-
ER Aur	-	-	West-14	West-14	Draco2	Draco2	-	Draco2	Draco2	Pt5n
BK Aur	-	-	West-14	West-14	East-16	Draco2	-	Draco2	Draco2	-

02/03/23	04/02/23	06/03/23	07/03/23	08/03/23	09/03/23	10/03/23
-	-	-	West-14	-	-	-
-	-	-	-	-	-	-
-	-	-	-	-	-	-
-	-	-	-	-	-	-
-	-	-	-	-	-	-
-	-	-	Draco2	-	-	-
-	-	-	West-14	-	-	-
-	-	-	West-14	-	-	-
-	Pt5m	-	-	-	-	-
-	Pt5m	-	Draco2	-	-	-
-	Pt5m	-	-	-	-	-
Pt5m	-	-	-	-	-	-
-	-	-	Draco2	-	-	-
Pt5m	-	Pt5m	-	-	Pt5m	-
Pt5m	-	Pt5m	-	-	Pt5m	-
-	-	-	Draco2	Pt5m	-	-
Pt5m	-	-	-	-	-	Pt5m
Pt5m	-	Pt5m	-	-	-	-
-	-	-	Draco2	-	-	-
Pt5n	-	Pt5m	-	-	Pt5m	-
-	-	-	Draco2	-	-	-

Scientific Summary for a General Audience

Cepheid variables are stars which change their brightness in a pattern over a period of time. Each Cepheid has their own pattern that they follow which occurs over a different time period for each. By measuring the apparent brightness of the stars, the pattern that their apparent brightness follows was mapped out allowing for the time it takes for this to complete one cycle to be calculated. Each Cepheid will have a different brightness and just because we see some of the Cepheids appear brighter than others on Earth it does not necessarily mean that they are actually releasing more light because the Cepheids are each at different distances from Earth and so if there were two Cepheids of the same apparent brightness (and so releasing the same amount of light) at different distances from Earth, for example, the Cepheid that was closer to Earth would look brighter to us.

Using the distance to each of the Cepheids, alongside the average apparent brightness throughout each cycle, a relationship was developed between the time it takes to complete a cycle (period) and the amount of light released (brightness), showing that as the period of the Cepheid increased, the star got brighter.