

Chapter 3

Data acquisition, reduction and photometry

3.1 Introduction

This chapter will be used to explain the steps and processes involved in doing the optical observations of the two open clusters NGC 6204 and Hogg 22, together with the data reduction and photometry.

In the current era, data acquisition relies quite strongly on the properties of the CCD (Charge-coupled device) detector. Prior to the use of CCDs, photo multiplier tubes (PMTs) and even earlier photographic plates were used to record astronomical data. The data obtained from the CCD is in a raw format, which needs to be reduced in order to improve the accuracy of the subsequent photometry process.

The data reduction that will be explained in this chapter involves flatfielding, debiasing, dark subtraction and image rotation. These tasks were carried out on all the obtained image frames.

By this being a variability study, calibration of the photometric results to a standard magnitude system was not necessary because differential photometry was performed on the cluster stars. However, in order to compare the data with previous observations of these two clusters, calibration of the observed instrumental magnitudes followed the process of photometry. Photometry was performed by using the IRAF software package, in combination with software developed to deal with the rotations and translations in the images obtained from the CCD camera.

A significant part of this study involved the development of software which can be used to analyse cluster data, doing photometry on all the stars in the field and extracting variable stars.

The two clusters display a sparsely populated region, in which around 3000 stars could be identified in one frame using the IRAF DAOPHOT task. During the 13 nights, a total of 2164 frames were taken, which highlights the importance of automation of the process of data analysis as a whole.

3.2 Data acquisition

3.2.1 Setup and instrumentation

- **Telescope**

The small observatory of the North-West University, South Africa was used to obtain the data for this study. In early 2010, a new telescope was assembled at the observatory, replacing the old 12 inch Meade LX200 telescope.

The observatory site is located on the Nooitgedacht farm, $-26^{\circ}54'09.34''\text{S}$, $27^{\circ}10'50.33''\text{E}$, at an elevation of 1448 m, and approximately 35km from Potchefstroom, South Africa.

The new Meade LX200 telescope with a 16 inch aperture, 4064mm focal length and f/10 Schmidt-Cassegrain optical tube assembly is mounted on a German equatorial (see Figure 3.1), computerized Paramount ME mounting system which is controlled through *The Sky 6* software package. The mounting system of any telescope is an important, if not the most important, aspect of the whole setup because of the stability and precision in drive which it must provide to the optical system.

- **CCD camera**

The main CCD camera used in this study was a QSI (Quantum Scientific Imaging) 540wsg, with a 2048×2048 pixel, interline transfer, progressive scan CCD detector. The pixels are $7.4 \mu\text{m}$ square pixels with microlenses and a full well capacity of 40 000 e.

For open cluster work it is advantageous to have a large field of view (FOV).



FIGURE 3.1: The new 16 inch telescope setup on the small observatory, Nooitgedacht, of the North-West University, where the more recently installed autoguiding system can also be seen.

The FOV depends on both the optical properties of the telescope and the properties of the CCD detector. This can be calculated for a specific detector by first computing the plate scale in arcseconds per pixel and then multiplying the plate scale by the size of the pixel array. The plate scale P is given by $P = \frac{206265 \times \mu}{1000 \times f}$, where μ is the pixel size in microns and f is the focal length of the primary mirror in mm (Howell, 2006). With a plate scale of 0.37 arcseconds per pixel, this telescope and CCD camera combination yields a 12.8×12.8 arcminute field of view.

An internal filter wheel with Johnson U, B, V, R, I filter system was installed, however only B, V and I filters were used during the observations of the two open clusters. During the observations, the CCD sensor was cooled internally to -20° C, in order to reduce the dark current generated during operation.

3.2.2 Observational strategy

The task of acquiring astronomical data, is highly influenced by the type of study conducted. Careful attention should be paid to creating an observational strategy in order

to succeed in reaching the final objectives of a particular study.

The observation of NGC 6204 and Hogg 22 was done over a time period of 13 nights between 27 July 2010 and 21 August 2010. With a variability study like this, it is necessary to have as many data points as possible in order to be able to observe both short and long term variability. In this study the Johnson V band was mainly used with occasional filter changes to Johnson B and I band.

The typical exposure times that should be used, depend mainly on the aperture size of the telescope and the magnitude of the target object. In the case of NGC 6204 and Hogg 22, cluster members range in brightness between 7th to 14th magnitude (Forbes & Short, 1996). An exposure time which resulted in the two or three brightest stars in the field being overexposed was chosen for this study. For the B, V and I bands, the exposure times was around 70, 60 and 40 seconds respectively, where the exposure times were also occasionally altered due to an increase or decrease in the atmospheric seeing conditions which ranged between 1.5 and 2.4 arcseconds.

Evening and morning flatfields, around ten frames in total, were taken on each observing night with a mean level between 20,000 and 30,000 analog to digital units (ADU). During the observing nights, a few bias frames were also taken in order to check on the change in the bias level of the CCD detector. It is also best to take bias frames at a stabilized temperature, which is at the temperature that observing takes place during the night.

After each observing night, a series of dark frames were taken with the same exposure time as the science images, in order to correct for the dark current in the detector (private communication).

3.3 Data reduction

Data reduction involves all the steps between observations and the start of photometry on the image frames. Raw CCD frames contain information on the observed object together with noise. Noise can be classified as random noise or systematic noise. Random noise can be accurately measured and described statistically but cannot be eliminated while known sources of systematic noise can be eliminated from and corrected for in the CCD frames. Systematic noise can be subdivided into additive noise and multiplicative noise, where bias and dark corrections belong to the additive class and flat corrections to the multiplicative class (Birney et al., 2006).

Other than systematic noise removal, image reduction also involves any physical change being done on the image frames. In this study, an additional step involving the rotation

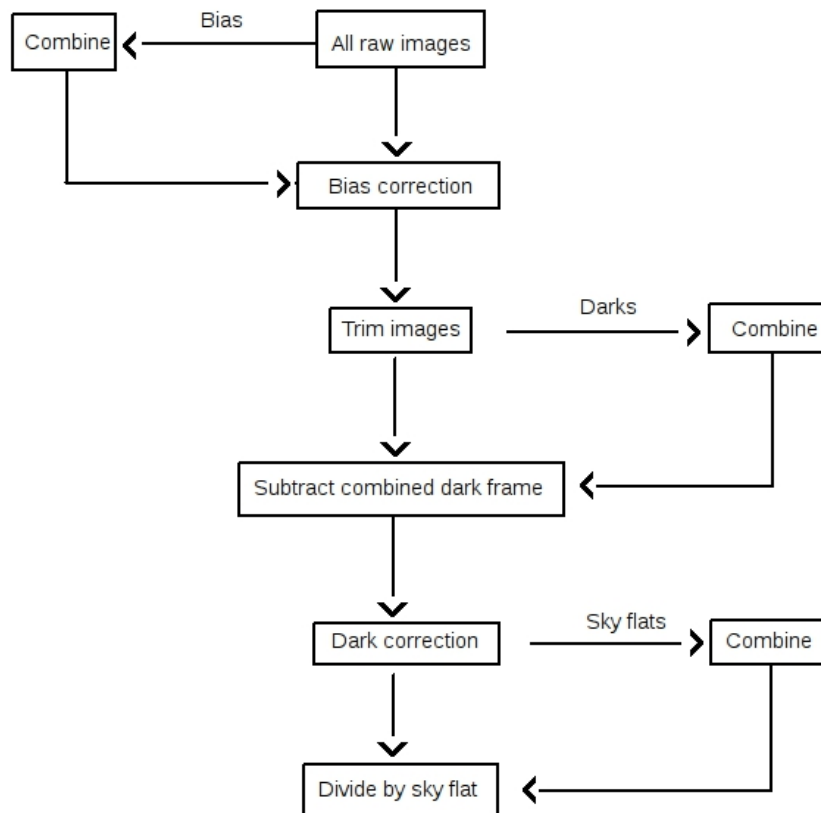


FIGURE 3.2: Schematic diagram of reduction process followed in this study

of a number of image frames was done in order to obtain the same orientation throughout the data set. As a result of the German equatorial mount of the telescope, all image frames obtained East of the meridian was rotated through 180 degrees compared to the images taken West of the meridian. This occurred because of the rotation of the optical tube, called the meridian flip manoeuvre, which is executed when the mount slews into the meridian from the east side and causes the optical tube to reposition so that the orientation of the telescope is not inverted when observing to the west side of the meridian. The first step in the reduction process involved the initial rotation of all images taken on the west side of the meridian through 180 degrees in order to work with images with similar orientation. The instance on each night where the meridian flip took place, was located in the image files and then the initial rotation was done by using the IMROT task of the WCSTOOLS package.

A flow chart of the steps involved in the data reduction process that will be explained in more detail in the rest of this section can be seen in Figure 3.2

3.3.1 Debiasing

Bias frames are CCD images with an exposure time of zero seconds during which the shutter does not open and the CCD is read out. The bias image allows the observer to determine the underlying noise level within each CCD frame, which originates from the CCD on-chip amplifiers (Howell, 2006). The purpose of the bias level, which is added to the signal, is to prevent negative readout values to arrive at the analogue to digital converter (ADC). Owing to this offset to the signal, a situation where no signal is recorded will always correspond to some positive bias signal (McLean, 1997).

Between eight and sixteen bias frames were taken each night depending on the duration of observation. The frames were then average combined using the IRAF zerocombine task located in *noao.imred.ccdred*. To debias each night's data, the average bias frame was subtracted from all image frames including dark frames and flat fields.

In order to analyse a typical bias frame, one of the average combined bias frames was used to create a histogram of the number of pixels as a function of the ADU value of each pixel. Figure 3.3 shows a histogram of the master bias frame which is the averaged frame from sixteen single bias frames, taken on a single night. The width of the ADU distribution shown by this histogram can be related to the gain and read noise of the CCD by the following relation:

$$\sigma_{\text{ADU}} = \frac{\text{readnoise}}{\text{gain}} \quad (3.1)$$

It can be seen from the histogram that σ_{ADU} is around 15, measured as the full width at half maximum (FWHM) of the distribution. The gain of the CCD is 0.75, resulting in a read noise of around 11 electrons (Howell, 2006).

3.3.2 Dark frames

The dark current in a CCD detector can be described as the current originating from thermal electrons freed from the valence band in the silicon from which the detector is constructed. These thermal electrons can then be collected in the potential well of a pixel and will be read out with the signal from astronomical photons. According to Howell (2006), this dark current is highly dependent on temperature where a large number of electrons can be expected to be read from a CCD detector at room temperature. However, typical dark current values of properly cooled CCD detectors can be statistically less than 1 electron/pixel/second and therefore by effectively cooling the CCD device, the resulting dark current can be rendered insignificant.

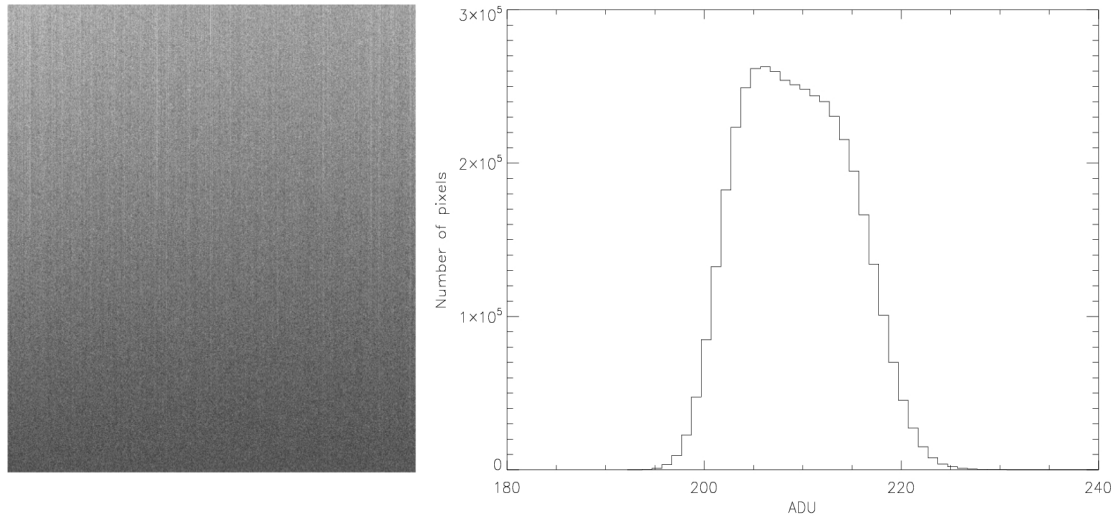


FIGURE 3.3: Left: Averaged bias frame. Right: Histogram of the average combined bias frame showing the number of pixels as a function of the ADU value of each pixel. The mean bias level is around 208 ADU.

Due to the fact that it was not possible to cool the specific CCD down to a minimum temperature of $-80^{\circ}\text{Celsius}$, the dark current could not be neglected and had to be corrected for (Birney et al., 2006).

In contrast to the bias frames, dark frames are taken with a non-zero exposure time, however the shutter also does not open during the exposure. Similar to the bias correction, the dark current can also be corrected for by subtracting an average dark frame from all the data frames taken during a night. A total of 10 dark frames, for each exposure time used in the science images, were taken during each night and average combined using the IRAF DARKCOMBINE task located in *noao.imred.ccdred*.

3.3.3 Flat fielding

Flat field frames can be used to correct for multiplicative noise found in CCD images in the form of quantum efficiency (QE) variations between pixels or non-uniform illumination. The three different sources of uniform illumination that can be used for taking flat fields are the twilight sky, an illuminated dome screen and the night sky, of which the twilight sky being the most accessible source. The night sky flat fields will not be described below, for it is mostly used in the infra-red wave band, which have a brighter sky background (Birney et al., 2006).

- **Twilight sky flats**

In order to take good quality twilight flat fields, thorough planning needs to be done to avoid missing the small window of opportunity to take sky flats due to the

fast changing sky brightness. Sky flats are known to contain low spatial frequency information or gradual variations across the image, which is due to the high uniformity of the twilight sky. The most suitable position in the sky to take sky flats is 20 degrees east of the zenith just after sunset and the opposite position of the zenith for morning sky flats. At this region in the sky, the brightness gradient is the smallest (Birney et al., 2006).

In the case of stars appearing in the sky flat frames, removal of the stars can be done either by moving the telescope between consecutive exposures or by disabling the tracking of the drive. When following this procedure before combination of the frames, the average frame will contain no stars or trails of stars. It can, however, happen that if the drive is disabled, star trails may overlap and might still be visible in the final combined image. Careful inspection of the final flat image is, therefore, necessary (Birney et al., 2006). During the observations of this study, the telescope drive was disabled during the flat field exposures and the final combined flat field image was inspected for any left over star trails after combination.

- **Dome flats**

A different type of flat field image can be obtained by taking an exposure of an illuminated screen inside the telescope dome called a dome flat field. One distinctive advantage of taking dome flats is that there is no urgency to get it done in a specific time like in the case of sky flats. However, dome screens are not as uniformly illuminated as the twilight sky. In spite of this drawback it is possible to combine the two types of flats to create a flat field image containing the low spatial frequency information of the sky flat and the high spatial frequency information of the dome flat which is more variable on a pixel scale (Birney et al., 2006).

3.4 Photometry

One of the major challenges during the course of this study involved the automation of the photometry process, where differential photometry had to be done on all stars in the obtained image frames. The photometry process followed in this study can be described by the following tasks:

- Coordinate transformation from a master image to the image under study

- Using the transformed coordinates to do point-spread-function (PSF) fitting and aperture photometry on all stars in the field
- Doing differential photometry on the output from the PSF and aperture photometry

When doing photometry with the IRAF Daophot package, a coordinate file with the coordinates of all stars is needed for each image. These coordinates correspond to the center of the object on which the photometry will be done. Due to the fact that the data set comprises of more than 2000 image frames, the option of creating a coordinate file for each image will not be an efficient method to follow, therefore the quest for automation.

3.4.1 Preparation for automated photometry

After the reduction process, inspection of the images revealed a translational movement in the images over the course of each night's observations. Movement that occurs in image data taken during an observing run will complicate the photometry process greatly in the sense that a coordinate file has to be created for each image. As a result of this the comparison of stellar magnitudes after photometry will be more difficult due to the fact that the same star will not have the same position or ID in multiple frames to identify it automatically. This movement that was seen in the images was due to unguided observing, when at the time of observation, no guider system was installed yet. This created the problem, when doing repeated photometry of all the image frames, that not all stars were present on each frame and do not have the same position on all frames.

Another problem involved the numbering of stars. The numbering is quite important in order to identify the stars in multiple images. To solve these two problems, a master list of stars was created which included the x-y frame coordinates of all stars on a single master frame. The frame that was chosen to be the master image, from which the master coordinate list was constructed, had to be a good image in the sense that it had to be taken during good atmospheric seeing conditions.

This master coordinate list was then sorted according to star brightness resulting in a numbering scheme which would be the same in each image. This was done by using the IRAF DAOFIND task (*noao.digiphot.daophot*) to find all the stars in the master image and using the PSORT task to sort them according to brightness. The master list was then created by extracting the sorted star coordinates from the IRAF Daofind output file.

Due to the fact that only one image was used to create the master list, not all stars were included in the list because of the movement between images. For this reason the

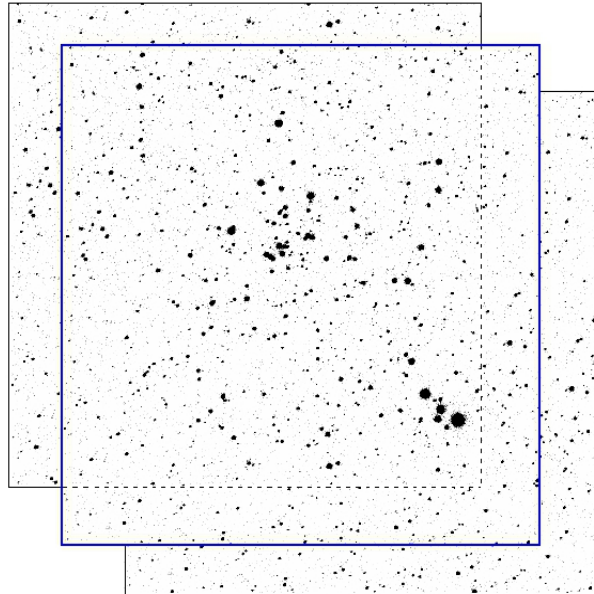


FIGURE 3.4: Construction of the master list of stars by using the most shifted images together with the master image (blue).

images which shifted most were also used to add stars to the master list as displayed in Figure 3.4. When this was done the effective region covered by the master coordinates was larger than any single image frame.

With a complete master list, it was then possible to do a coordinate transformation from the master list coordinates to a specific image, so that photometry could be done repeatedly on a list of images. In what follows, the image under study will refer to the image that was read from a list on which photometry was done of all stars.

A C-code was developed to do the coordinate transformation from the master list to the image coordinates. Before doing this transformation, a quick DAOFIND run is necessary to produce coordinates of the 30 brightest stars in the image under study. The coordinate transformation then relied on the calculation of distances between the 30 brightest stars found in the image under study and then comparing the distances to the distances between the brightest stars on the master image. The distances were then cross correlated in order to find the same stars in both the master image and the image under study.

After locating the same stars on both images, their coordinates were used to calculate the rotation and translation between the master image and the image under study. The rotation and translation were calculated by simultaneously solving the two coordinate

transformation equations given by,

$$\begin{aligned}x &= x' \cos(a) - y' \sin(a) + x_0 \\y &= x' \sin(a) + y' \cos(a) + y_0\end{aligned}\tag{3.2}$$

where x, y are the pixel coordinates in the master frame and x', y' denote the pixel coordinates from the image under study. The angle of rotation between the image under study and the master image is given by a , and x_0, y_0 are the translational distances in pixels from the master image coordinates. Finally the complete master list with 3082 coordinate entries was transformed to the reference frame of the image under study by substituting the calculated rotation angle and translational distance into equation 3.2.

This method then resulted in photometry being done on the same number of stars in all images where the number of stars are independent of the seeing conditions at which the image was taken. In the case of photometry being attempted on stars that do not appear in a specific image, which can be due to bad atmospheric seeing conditions, the resulting magnitude will not be added to the time series of the particular star and will not cause any problems with later analysis.

The next step in the photometry process involved the PSF and aperture photometry that was done by using the DAOPHOT IRAF package. An IRAF Photometry script, created by Eran Ofek (Ofek, 1997), was used as starting point for constructing a new IRAF photometry script suitable for use on this data set and similar data arrangements. The C code that carries out the coordinate transformation together with the differential photometry procedure that will follow had to be included in the script to ensure efficient automation of the complete photometry process.

3.4.2 PSF photometry

Due to the sparsely populated nature of NGC 6204 and Hogg 22, both PSF and aperture photometry could be done on this region. PSF photometry is predominantly used on more densely populated stellar regions while aperture photometry yields more acceptable photometric results performed on well separated stars (Howell, 2006). The process of PSF photometry involves the fitting of a predefined mathematical function to the light profile of the astronomical object recorded on the CCD. Some of the most common functions used in profile fitting include the Gaussian, Lorentzian and Moffat functions. A Gaussian function was used in the profile fitting which is defined in the IRAF DAOPARS

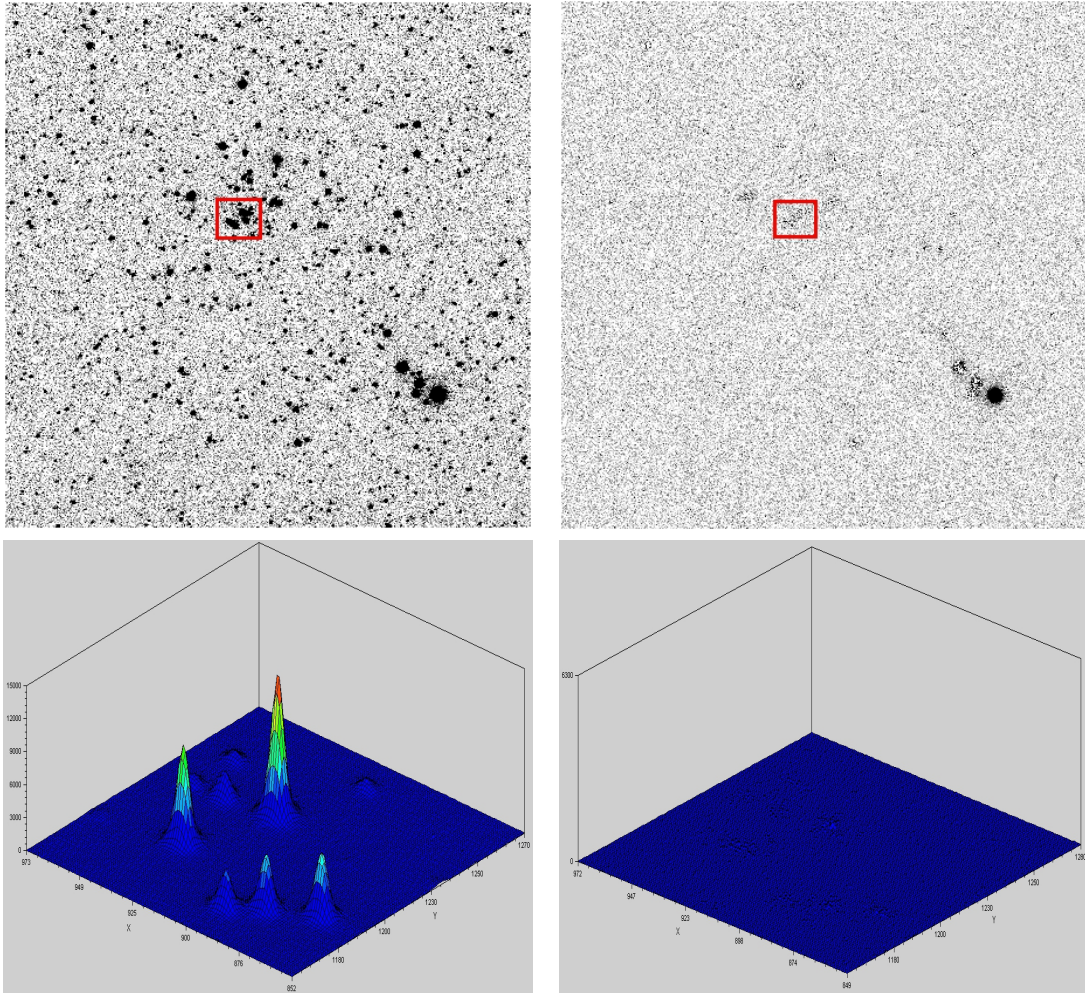


FIGURE 3.5: Field of NGC 6204 and Hogg 22 shown before (left) and after (right) subtracting the PSF function.

parameter file as

$$Ae^{(-0.5z)}, \quad (3.3)$$

where A is an amplitude or normalization constant and $z = x^2/p1^2 + y^2/p2^2$. $p1$ and $p2$ are parameters which are fitted during the generation of the PSF model (IRAF help function). This profile which is present in each point-like object recorded on the CCD, is a distinctive property resulting from the specific optical arrangement. An example of the light profile of stars from the data set can be seen in Figure 3.5.

In order to create a profile which is representative of most stars in the image, some stars must be selected as PSF stars, from which the PSF will be constructed. The selection of stars from the master image to be used as PSF stars requires thorough inspection in order to avoid contaminating the PSF model with inappropriate stars, which will affect photometry accuracy. The norm is to use as many unsaturated and well separated stars

with an acceptable light profile in order to obtain a PSF model with which little to no residuals are left on the image after subtraction of the model. In this study a total of 317 PSF stars were used and was only specified by star number from the master list of all stars to be able to use the same PSF stars in all the images. The PSF stars were selected by using the IRAF PSTSELECT task in the package *noao.digiphot.daophot*. When running the task, a number of selected PSF stars will be displayed where the user can accept or reject the star in question.

Before starting to construct the PSF from the selected PSF stars in the image, instrumental magnitudes of all stars are needed as a starting point for the PSF construction. This was obtained by running the aperture photometry task, PHOT, in *noao.digiphot.daophot* and using the resulting aperture photometry file as input to the PSF task.

A successful PSF fitting relies on a number of variables that need to be defined in the IRAF script or parameter file for the PSF to be an adequate fit to the light profile of the stars. The main parameters used are the full width at half maximum (fwhm), minimum and maximum good data value, PSF fitting radius, the varorder, which is a parameter set for the variability of the PSF function over the image, and the sky noise of the image under study.

PSF construction is an iterative process, where the PSF, NSTAR and SUBSTAR tasks are used in succession until a clean subtracted image is produced. The ALLSTAR task is then used to fit the PSF and measure all the stars found in the image. In this study, three PSF iterations were used together with a PSF function that varies quadratically over the image. This variability of the PSF can be set by changing the varorder parameter in the *daopars* parameter file. With a varorder of 2, as used in this study, the analytic function together with six look-up tables were used to compute the PSF model (IRAF help function).

The end result of the PSF photometry process will reflect in the subtracted image, produced from the SUBSTAR (after each iteration) or ALLSTAR (after final iteration) tasks. When the subtracted images were inspected and no residuals could be seen from the stars, the PSF was a good fit to the profile of the stars, as can be seen in Figure 3.5 where the first column shows the initial master image (top) and PSF of stars (bottom) and the second column, the final image after subtraction of the PSF model.

3.4.3 Aperture photometry

An aperture used as a digital area to measure the flux from a star, should be chosen carefully in order not to exclude the wings of the stellar light profile which extend much further than might be expected from visual inspection. The stellar profile is influenced by atmospheric refraction and diffraction and instrumental diffraction. The profile of a star consists of three parts, a nearly uniform centre disk, a Gaussian decrease and an exponential decrease which extends up to a factor of 1000 in radial direction (King, 1971). By plotting the stellar magnitude as function of aperture size, a good indication of the optimal aperture size can be obtained by looking at the region where the magnitude starts to level off into the image background level.

As stated earlier, aperture photometry done on bright isolated stars almost always surpass the accuracy of PSF photometry on these stars. By doing both aperture and PSF photometry, a comparison of PSF and aperture photometry can be done and the best photometry for all stars in the entire magnitude range can be obtained.

In this study, aperture photometry was performed using the IRAF PHOT task by using four different aperture sizes. The aperture sizes were scaled according to the atmospheric seeing conditions by using two parameters from the IRAF PSF fit. These parameters, found in the PSF output file (PAR1 and PAR2), are an indication of the atmospheric seeing conditions during observations where they indicate the FWHM in the x and y direction of the PSF profile. The sum of the FWHM in x and y were calculated and multiplied by 1, 1.5, 2 and 2.5, in order to obtain four different seeing scaled aperture values.

3.5 Differential photometry

Differential photometry involves comparing sources located in close proximity to each other in the sky, which also means that differential first-order extinction can be neglected (Birney et al., 2006). One major advantage over all-sky photometry is the fact that changing atmospheric and seeing conditions, during an observational run, affect all stars equally and do not negatively influence the success of differential photometry.

When applying this technique to a stellar field, comparison stars should be located in the same field which tends to be constant over the observational period. Differential photometry was done on this region by using the average magnitude from five bright stars in the field, which showed to be the least variable over the observed time period.

These "constant" stars were selected by visually inspecting their light curves (Figure 3.6) in order to identify the stars with the lowest amplitude variations. In a variability study

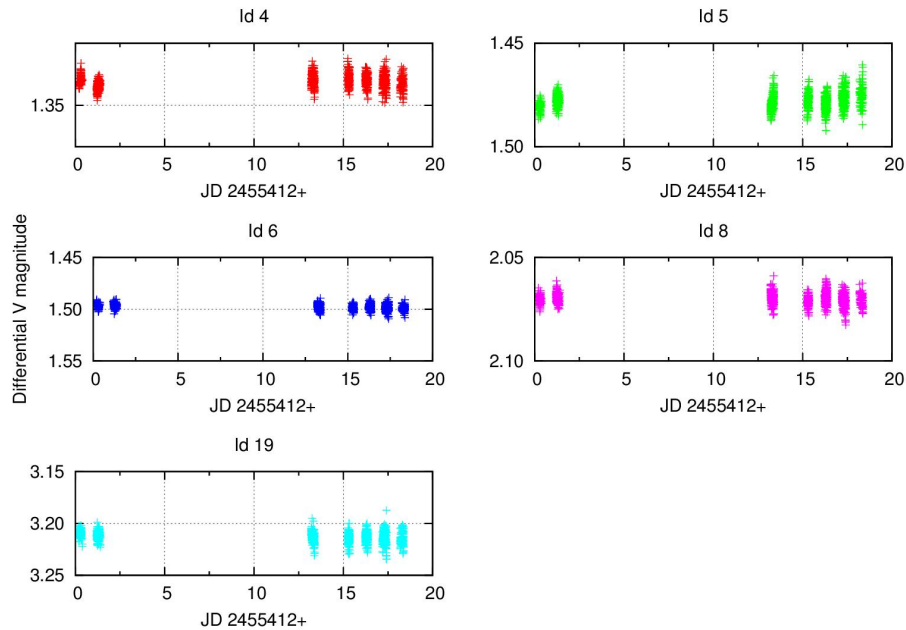


FIGURE 3.6: Light curves of the stars selected to be used as constant stars in the differential photometry process. Note that the differential V magnitude is indicated on different scales.

like this where periods are extracted from variable stars in the clusters, transformation of the observed magnitudes to a standard system is not necessary. However, when the photometric results are compared to results from the literature, the transformation process is of great importance and cannot be neglected.

3.5.1 Cleaning of time series data

An important step before attempting to extract variability from the data is to clean the data set from outliers which will influence the results obtained from the application of the Lomb Scargle transform. These outliers can occur as a result of many different circumstances, ranging from atmospheric conditions to instrumental effects.

The process of cleaning is overseen by the choice that will determine the amount of outliers to be excluded from the data set. This was done by performing a four sigma clipping on the data, where the average of all data points from each light curve was used as zero level from which four standard deviations were added to create the upper limit and subtracted to create the lower limit. This was performed only on the magnitude values, where the errors on the magnitudes were calculated in the differential photometry code supplied in Appendix B. All the data points between these two boundaries

were accepted to be included in the final data set on which the Lomb Scargle transform was applied. Figure 3.7 shows the light curve of one known Beta Cephei star observed over seven nights, located in the Hogg 22 cluster, together with the two lines indicating the four sigma clipping. One outlier can be seen in this case, from the second night, which indicates the importance of removing the outlying data point before applying the Lomb Scargle transform.

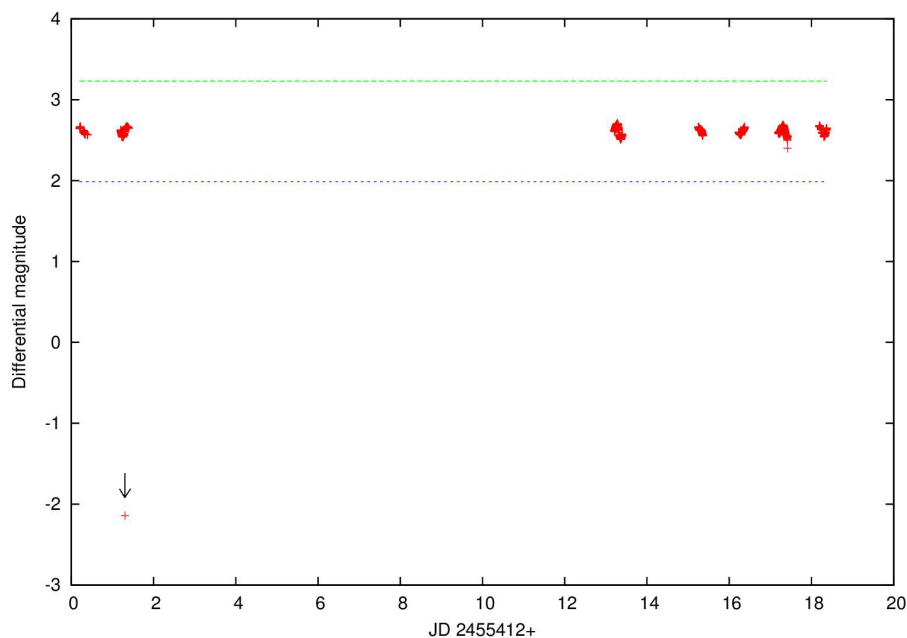


FIGURE 3.7: The light curve of the known Beta Cephei star from Hogg 22, shown under a four sigma clipping. One outlier can be seen in the figure from the second night of observation. The green line shows the four sigma level above the average of the light curve and the blue line indicates the four sigma level below the average.

The choice of severity of the cut by which data will be excluded is not a well defined one, but it will depend on the properties of the data set. In this case, however, a visual inspection of the light curve proved to be a very useful aid in selecting between real and spurious data points.

3.6 Timing

Timing in astrophysical studies, in particular variability surveys is of great importance because of the relative motion and position of the Earth and the position and time at which the observations were done. In order to compare observations of a source done over extended periods during a year, a timing standard needs to be established. As the Earth moves around the Sun, the difference in arrival time of light can differ with as much as 16^m and 55^s , when an object is observed from opposing sides of the orbit.

For this reason the Heliocentric Julian date (HJD) needs to be used in order to correct the time to an observed time as from the barycentre of the solar system (Birney et al., 2006). In the current study, the conversion from Julian date (JD) to HJD was done by using three different conversion codes (IRAF, C-code and Web based) where the output from all three were compared. By comparison of the results it was found that none of the conversions agreed where a difference of 63 seconds were found between the IRAF conversion and the C-code. After correspondence with Prof A. Pigulski, it was decided to use only the unconverted JD in this study.

In the following section of results and discussion a detailed explanation of different aspects relating to the output from the data reduction and photometry pipeline will be provided as well as the success of its application to the data set. Questions on the choice of an optimal variability level used to define the lower limit of the variable sources, extracted with the Lomb Scargle transform, will be addressed together with possible classification of high amplitude variable sources. After calibration of the instrumental magnitudes to a standard system, photometric results can be compared to results from previous studies done on these two clusters.