Chapter 2

Observations and Data Reduction

2.1 Introduction

With the main thesis aims chosen, a careful choice of telescope and detector had to be made in order to best meet those aims. With 47 Tuc as the set target, which has a known radial extent, the choice of telescope and detector is important in order to sample as much of the cluster as possible in a single exposure. Successful transit observations require not only a large sample of stars, but also high temporal resolution. This is in order to have a large number of data points sampling the whole of the transit duration. In summary, a transit search requires a wide field of view plus a sufficient aperture to allow for short exposure times.

As 47 Tucanae is a southern object (Dec=−71.1°), in order to be able to observe it for a long contiguous window, the observing site must be located at a suitable latitude. Siding Spring Observatory (SSO, a part of the Mount Stromlo and Siding Spring Observatories, MSSSO) was chosen. This site is at a latitude of −31.2°, and as such presents an excellent location to observe the cluster. As well as location, the median seeing of the site is suitable. At SSO, this has been measured previously at 2" (Coleman et al. 2004) Full Width Half Maximum (FWHM), with average levels of cloudy nights for Australia (at about 50%). As derived from the AAT WFI exposure time calculator\(^1\) and scaled down to the aperture of the 40", the seeing for this project must be better than 2.5" to successfully observe a transit at a signal-to-noise level sufficient for a 3σ detection. As a side-point, SSO is also the site of the UNSW APT planet transit search (Hidas et al. 2003).

Located at SSO is the Australian National University 40" (1m) telescope. This telescope was the first to be placed on the site, and is generally available for long observing runs. Periodically attached to the 40" is the Mount Stromlo Wide Field Imager (WFI), which provides a significant field of view of 52'×52'. This combination of telescope and detector is perfectly suited to searching for transits.

\(^1\)http://www.aao.gov.au/cgi-bin/pfcalc.pl
This chapter describes the hardware used to obtain the photometric dataset, and how this is useful for the thesis aims. Before the main observing proposal was submitted, Monte Carlo simulations were performed by P.D. Sackett to derive the expected numbers of transits that should be observable within a dataset obtained in average SSO observing conditions, in order to check if the project was technically and statistically feasible. These dataset expectations are presented and discussed. The details of the actual observations as carried out at the telescope are presented, and the observing strategy outlined. All observations were obtained in one run, with a single associated strategy. By observing only one target, the length of continuous observation can be maximized, leading to a dataset that is useful for many research opportunities.

After the observing run, the dataset must be reduced to produce suitable research images. The reduction techniques remove as much as possible the effects of the atmosphere and systematic defects inherent to the telescope/detector system. The result of this is a dataset which is internally consistent, and suitable for the main photometric analysis. As planetary transits are of small amplitudes, the photometric reduction is critical to meet the thesis aims; this is explained in this chapter, with the necessity of each step described.

A description of the final dataset after completing this reduction process, including a discussion of the ability of the data to contain visible transits is then described. As a part of this run, data were taken in two bands (V and I) in order to produce a Colour Magnitude Diagram (CMD) dataset for 47 Tuc. This is needed in order to place any detected transiting system or variable star onto the standard magnitude system. These data are presented, and the CMD discussed to introduce terms that are referred to in the rest of the thesis. A description of the CMD calibration process is also presented.

An astrometric solution was obtained for all stars visible on the best seeing image obtained during the run. This is discussed, along with the accuracy of the solution. This chapter leads into a detailed description of the way in which the photometry was derived, and also the details of the periodogram program used to search the dataset for variable stars.

2.1.1 Dataset Aims

The aim is for the main dataset to be able to provide time-series of sufficient quality to routinely detect a 0.01-0.02 magnitude signal for a large enough sample of stars, that a meaningful statistical conclusion can be made. This is of course dependent on the observing conditions, but the dataset quality can be maximised via the use of specific observing techniques.

The seeing of the data is most important for this, and a firm upper limit was
kept over the course of the run. By examining each image in turn manually for quality, this was easily accomplished, with the added bonus of immediately identifying images that contained satellite trails and other transient phenomenon. Also, telescope offsets were minimized, both from one night to the next and in-between images during the course of each night. This reduces the number of stars that are lost off the edges of the CCDs, and also helps to minimize spurious detections caused by any large intra-pixel sensitivity differences and CCD blemishes such as bad columns. Focus is important for maintaining image quality. Generally, once calibrated, the 40″ focus does not shift enough to produce measurable differences on the images. However, sudden changes in air temperature can cause shifts, and the focus was independently checked and adjusted after readout of each image.

2.2 Telescope and Detector

The magnitude range of the 47 Tuc stars to sample for transits is V=17.0-19.5, as these values correspond to main sequence targets. Considering the quantum efficiency of modern CCD detectors, a large aperture telescope is not required to reach these moderate magnitudes with sufficient photometric accuracy to detect a transit signature. To maximize the transit recoverability, this project required a great deal of contiguous observing time, as much as can be asked for realistically, accounting for average weather conditions. It follows that the longer the observing window, the longer the maximum detectable Hot Jupiter orbital period will be, and the higher the number of detectable planets that should be contained in the dataset.

As telescopes go, the 40″ when fitted with WFI is perfectly suited to the task at hand. Since the cluster has been estimated to be 23′ in diameter (Harris 1996), this telescope and detector combination is easily capable of observing a large portion of the extent of 47 Tuc in one field of view. The 40″ has Ritchey-Christien optics, and as the telescope is equatorially mounted, field rotation is not an issue when observing for extended periods of time.

When attached to the 40″, WFI provides a 52′×52′ field of view (diagonal approx 1.2°), which is easily capable of encompassing 47 Tuc in a single exposure. The detector consists of a 4×2 array of 2048×4096 pixel back-illuminated CCDs. These CCDs are arranged to give a total format of 8k×8k 15 micron pixels. The CCD layout of WFI is shown in Fig. 2.1, an image taken from the Mount Stromlo WFI website\(^2\). The detector yields a scale of 0.38″/pix at the 40″ Cassegrain focus. This pixel scale is much smaller than the median seeing of the site, allowing a stellar image to be spread over multiple pixels. The CCDs are cooled to 180K via liquid nitrogen refrigeration, reducing the thermal current. The readout time for

\(^2\)http://msowww.anu.edu.au/observing/wfi/
Figure 2.1: The CCD layout of WFI. The eight CCDs making up this detector can be seen. This detector provides a 52’×52’ field of view (diagonal approx 1.2°), and is perfectly suited for use in transit searches. From bottom right to top right (clockwise), the CCDs are numbered 1→8. When mounted on the 40”, north is to the right and east is down.

WFI is approximately 40 seconds, which is extremely fast considering the 64Mpixels that make up the array. With this fast readout, the ‘dead-time’ (telescope time that is not available for data-taking) can be reduced. Over a long run, this fast readout time makes a huge difference to the final size of the dataset.

When discussing CCD efficiencies, two other effects must be considered; readout noise and gain. Readout noise is defined as the number of electrons introduced into the final signal by the readout of the device. The lower the value of the readout noise the better will be the performance of the detector. As an example, if the readout noise is ten electrons, then this number of electrons will be added to the output signal per pixel. The gain of a CCD camera is defined as the conversion factor between the number of electrons (‘e-’) recorded by the CCD and the number of digital units (‘counts’) contained in the CCD image pixels. It is useful to know this conversion for evaluating the performance of the CCD camera. Since quantities in the CCD image can only be measured in units of counts, knowing the gain permits the calculation of quantities such as readout noise and full well capacity in these fundamental units. The units of gain are in electrons per count. Typical values of the readout and gain for WFI are 3.5-5.5 electrons and 1.5-2.1
electron per Astronomical Data Unit (ADU) and are dependent on the thermal current.

### 2.2.1 Filter and Exposure Times

In order to keep the exposure times to an acceptably small length, hence maximizing the in-transit sampling and while keeping a high signal-to-noise for stars of $V=17–19.5$, a single broadband Cousins V+R filter was employed. Specifically built for this project at SSO, this filter permits a significant increase in signal-to-noise per set exposure time, while maintaining the image degradation due to atmospheric dispersion to an acceptable level. Estimated from knowledge of the throughput of the V and R passbands, in 2" seeing the photon noise signal-to-noise ratio per data point ($snr_i$) is typically 220 with a 7-day moon for a star of $V=18.5$. This reduces to 165 at times of bright moon.

The following relationship holds between signal-to-noise and standard magnitude error for making these calculations:

$$\sigma = -2.5 \log_{10} \left( 1 - \left( \frac{1}{SNR} \right) \right)$$  \hspace{1cm} (2.1)

where $\sigma$ is the required photometric accuracy in magnitude units and SNR is the associated signal-to-noise ratio. Using this relationship, the V+R filter provides a $\Delta$ magnitude error or 0.005 for 7-day moon and 0.006 for bright moon conditions, with the associated signal-to-noise values previously presented.

The depth of the HD209458 transit is $\sim 1.5\%$, or 0.015 mag (Charbonneau et al. 2000; Henry et al. 2000), requiring a signal-to-noise of 200 or more for a 3$\sigma$ detection. Integrations no longer than 300s are needed to sample the transit adequately, that is, to have 15 points across the transit and obtain a signal-to-noise of $\sim 775$ ($snr_i \times \sqrt{n}$) over a typical transit duration.

Due to 47 Tuc being intrinsically a very crowded object, it is expected that transits over a significant portion of the cluster core will be unobservable in the data. To keep the signal-to-noise ratio high elsewhere, the most crowded parts of the cluster must be allowed to saturate. As such, the inner 6' is not sampled in this work. This is however not a significant problem. The experiment is designed to be most sensitive to the outer halo of the cluster, and even without the saturated core, the total number of stars to analyse for transits allows a result with reasonable statistical significance.
2.3 Expected Results

Before the observing run, the expected number of detectable transits in a modelled dataset taken in average observing conditions was determined, in order to test if the project was statistically feasible. The preliminary calculations assume that the frequency of 47 Tuc Hot Jupiter planets is the same as that observed in the solar neighbourhood. This is the base assumption that is to be tested in this project. The number of planets expected to be found in the dataset, as presented in the original telescope proposal, was calculated by taking into account the following:

- The number of appropriate stars in the cluster, estimated to be 16,400 stars from turnoff at V=17.1 to V=19 (an initial assumption of the magnitude limit of the search) within an area 32' on a side, excluding the inner 5' radius.

- The present detection rate of Hot Jupiter planets by radial velocity techniques in the solar neighbourhood (2-3%). This was the preliminary assumption, which turned out to be a little high.

- The chance that such a planet will transit (about 7%).

- The transit duration (~2.5 hours depending on inclination, period and stellar type).

- Effects of stellar metallicity on transit depth and duration, with lower metallicity stars having slightly smaller radii.

- The most typical orbital periods (3-5d) of Hot Jupiters, averaged over their orbital phase, taken from radial velocity results in the solar neighbourhood.

- Observing duty cycle (typically 1/3 for single telescope).

- Realistic weather conditions (estimated to be 50% usable.)

By changing the length of the observing window (by trying to account for the unknown future weather conditions), the number of expected planets was determined. Using Monte Carlo simulations, the number of planets expected to cause detectable transits on the 40'' was thus estimated to be 5.6±2.4 with a 21-night observing window. For a 14 and 28-day experiment, the numbers determined were 2.6±1.6 and 9.2±3.0 Hot Jupiter respectively. Therefore, Poisson statistics suggest that a 21-day SSO null result will be significant to the 2σ level, and a 28-day experiment to the 3σ level. A 28-day dataset would also allow sensitivity to find 2.4 Hot Jupiters in which three-in-a-row transits are visible in the dataset. Any object observed to execute three contiguous transit events is the easiest to confirm and determine a period for with small uncertainty.
2.4 Observations

The globular cluster 47 Tucanae was observed with the 40″ and the WFI for a total of 30.4 contiguous nights from 22\textsuperscript{nd} August to 24\textsuperscript{th} September 2002. The chosen field centre was RA=00\textdegree:24\textquoterem:05.2\textsec, DEC=−72\textdegree:04\textarcminute:51.0\textsec, positioning the core of the cluster directly onto the centre of the WFI field of view. Table 2.1 shows the positions of the field centres for all eight CCDs of the WFI used during the course of the observing run. Fig. 2.2 shows the location of the cluster in the WFI field of view. This image is the one used as a master pointing image, and each subsequent image was shifted to match this one, minimising the telescope offsets throughout the dataset. The spatial coverage of the cluster is clear, as is the extent of the saturation in the inner cluster core.

Of this total observing time, \(~80\%\) was useful for the main planetary transit project, with a calculated mean seeing of 2.12″. The seeing distribution of the dataset as a whole is presented in Fig. 2.3. The mean seeing of the distribution can be seen as a dotted line, and is lower than the \(\leq 2.5″\) seeing required for a 3\(\sigma\) detection of a Hot Jupiter transit, as determined earlier. Therefore the atmospheric conditions were suitable for the main thesis aims to be met.

This seeing was measured directly from each image as it completed readout from the array while at the telescope. Using MSCEXAM within IRAF\textsuperscript{3}, the full width half maximum (FWHM) of the stellar point spread function (PSF) was measured for a randomly selected sample of stars. As this unit was measured in pixels, by using the pixel scale for WFI (0.38″/pix) this unit was converted

---

\textsuperscript{3}IRAF is distributed by National Optical Observatories, which is operated by the Association of Universities for Research in Astronomy, Inc., under cooperative agreement with the National Science Foundation.
Figure 2.2: The location of 47 Tuc in the WFI field. The ability of the detector to contain the majority of the cluster in one single field is apparent, along with the position of the cluster core on the data. This is the master field with which all the other images were compared for pointing accuracy.

to arcseconds. By using this method, any images that were contaminated by transient artifacts such as satellite trails and sudden cloud were discarded and immediately retaken at the telescope. Similarly, any image that had an unsuitable seeing measurement was discarded, and the upper limit for this was chosen to be 3.0". Such images were still useful for increasing the temporal coverage of the transit search, and the seeing limit defined above does not apply to variable stars.

The way in which the seeing is distributed among the actual Modified Julian Dates (MJD) of the dataset is shown in Fig. 2.4. This figure also shows the temporal coverage of the cluster, along with the gaps caused by bad weather and daylight. It can be seen that the number of points that lie below the 2.5" limit is much higher than the number of points above it. Nights of good weather are clearly seen, recognisable as those with little or no scatter in image quality over the course of the night.

Telescope offsets were minimised at much as possible while at the telescope. At the start of the night, a master field (see Fig. 2.2) was used to position the telescope to exactly the same pointing. The numbers of stars lost off the edges of the CCDs were reduced by performing this quick procedure. It was found during
Figure 2.3: The seeing distribution of the 47 Tuc dataset. The mean of the distribution was measured as 2.12'' and is marked as a dotted line. The upper limit on the seeing needed for a 3σ Hot Jupiter transit detection (at 2.5'') is plotted as a solid line. The majority of the dataset is better than this limit, indicating that the quality of the observing conditions was suitable for the main thesis aims.

the course of the night that the 40'' had a systematic declination drift at this position in the sky. Once this reached a level of a few arcseconds, the telescope was moved back to this same master field.

2.4.1 Dataset Properties

After the run was completed, a total of 1220 object images were in hand. (Additional to this were a few hundred reduction frames, comprising bias frames, twilight flat field images and darkframes). Therefore each time-series that results from this dataset will have 1220 independent points. If three transits are observed in any particular time-series, this large number of total data points would allow around 45 data points to be during a transit.
Figure 2.4: The seeing of the dataset as a whole plotted against MJD of observation. The 2.5" seeing limit for the transit search is plotted as a dotted line. The temporal coverage of the cluster is apparent, with the gaps caused by bad weather and daylight. It can be seen that the vast majority of the data points were taken in better than 2.5" seeing.

2.5 Image Reduction

Once the images had been archived, it was necessary to remove as many systematic effects as possible. This was done via an image reduction process. Initial reduction of the raw images was undertaken with standard practices within the MSCRED package of IRAF. MSCRED is a mosaic version of the CCDRED package, allowing for all eight CCDs to be reduced in one go. This permits a significant computing-time decrease for this procedure. As hundreds of reduction images were obtained over the 30.4-night observing run, they allow for accurate correction of time-dependant variations across the WFI array. The reduction images comprise bias frames, twilight flatfields, dark current frames and a dark sky flatfield. The necessity for obtaining these frames and their application to the dataset are described here.

As CCDs are subject to noise introduced by pixel-to-pixel variations as well as natural occurrences such as cosmic rays, it is necessary to remove as many of these effects as possible before the images are used. Bad columns can be seen in some of the WFI CCDs in Fig. 2.2. These are caused by hot pixels, i.e,
those with permanent high values. As the image is read out, the preceding pixels have their pixel values replaced by that of the hot pixel, and the end result is a line of bad data. These have an adverse effect on photometric quality; if a star temporarily drifts onto a bad column, it can mimic a transit signal. By minimising the telescope offsets through the run, as described earlier, this effect was also minimised.

The first of the reduction frames utilised were the bias frames. These are defined as dark exposures with the minimum possible duration, that is, exposures of zero length taken with the shutter closed. Once obtained, these frames allow a subtraction of the pixel values produced purely by reading out the CCD. This is known as the bias level, and contains the readout noise information, which is reduced by cooling the CCD chips via liquid nitrogen refrigeration.

The second type of reduction frame produced were so-called ‘flat field’ images. Flatfields are needed as each pixel in the array has a somewhat different quantum efficiency. This is due to small structural variations in the CCD (structure that might be present due to chip manufacturing processes) and filters or other elements, including dust on the filter and CCD surfaces. The flatfield measures the resultant pixel-to-pixel variations in sensitivity, so that they can be removed from the object images. If flatfielded properly, an image will have a negligible difference in background counts level from one CCD to the next — it will be, in essence, ‘flat’.

Twilight flatfields are important for removing the large-scale variability across the array caused by different pixel sensitivity levels, and also for removing the dust particles and coating residues that are overlaid on all images taken with a specific filter. These sensitivity effects can be very large, up to 10-20% across the array. These sensitivity effects must be removed as much as possible, in order to derive accurate photometry across the array.

Flats are taken by observing a uniform light source, usually the twilight sky. For WFI's wide field of view, it is very important to make sure that the twilight observed is as clear as possible. If clouds are present, the resultant flatfield will not be suitable, as the cloud can have different density over the 52' × 52' field. Every effort should be made to keep the moon out of the immediate area used for flatfielding. Any extraneous light entering the telescope will adversely affect the quality of the images. The method used to produce useable flats for this project involved taking exposures of between 6 and 90 seconds duration, and measuring the counts level of the image upon readout. The level should be maintained at 10,000 to 25,000 counts to reduce the effect of non-linearity on the chips. The exposure time is altered from one image to the next to keep the flatfield inside these parameters as the sky changes brightness.

The twilight sky is brighter at longer wavelengths, and so if multiple filters are
used during the night, then the order in which the flatfields are taken (short to long wavelengths) can maximize the time interval available for data taking. The images were offset between frames in order to remove stellar residuals from the median-combined master flat.

It is good practice to take twilight flatfield images both in the evening and in the morning, as the flatfield can change during the course of the night due to mobile dust particles on the filter and chip sensitivity level variations. One master flatfield is produced per filter, and then used to reduce the appropriate images. If conditions are bad, flatfields can also be taken by observing a uniform white object inside the dome that is illuminated by a lamp. These ‘dome-flats’ are generally thought to be inferior to twilight flats when taken with the 40″ telescope, and therefore were not used at all during the run.

Finally, dark frames are exposures of a length equal to the object image exposures, but taken with the shutter closed. Any extraneous light leakage caused by guide CCDs and the like can be removed from the final reduced images. For WFI, this light leakage is significant for the four outer CCDs, due entirely to the small guide CCDs, which lie outside the main array.

As many reduction images were obtained at regular intervals during the run, a batch covering a night were median combined within IRAF to produce a master bias, flatfield and dark frame. This procedure was used to produce a night’s worth
of data, with the next night’s flatfield and bias frame being separately produced to help remove any variations from night to night. A sample bias frame, twilight flatfield and dark frame used in the reduction procedure for the V+R filter are presented in Figs. 2.5, 2.6 and 2.7.

For completeness, a dark-sky flatfield was obtained, by taking 300s exposures of an uncrowded region of sky (the same exposure time as used for the object images), offsetting an arcminute or so between images. These frames were then median combined to produce a frame containing further information of the small scale sensitivity variations across the array. Tests indicated that large-scale image quality was degraded when using this flatfield, and it was decided to not incorporate the dark-flat into the reduction procedure.

This reduction was performed in an automatic procedure for the whole dataset. After checking the final reduced images for flatness (using IMSTAT) and pixel-to-pixel variations, the resultant 1220 object images were then suitable for the main photometric pipeline and subsequent analysis, which is the subject of the next chapter.
2.6 Cluster Colour Magnitude Diagram

As well as the main object images obtained with the V+R filter, a number of images were also taken of 47 Tuc in both V and I filters in order to produce a Colour Magnitude Diagram (CMD) for the cluster. These diagrams are very important tools for analysing the stellar content of a star cluster or galaxy. Analogous to a Hertzsprung-Russell Diagram, a CMD displays the colour of a stellar population plotted against its magnitude (often in V). The evolution of the population can therefore be observed as the stars change both in brightness and colour as they evolve from the main sequence. Needed to obtain brightness and colour information for stars in the field of view, these data are vital for determining the nature of any transiting object and the likely type of the variable stars identified.

For production of the 47 Tuc CMD, a small number of images in both V and I were obtained with both 300s and 600s exposures, enabling a deep photometric catalogue of the stars to be made. The conditions when these data were taken were not optimal, being above the three arcsecond limit to the seeing needed for the transit search in an effort to make use of all conditions. Subsequent analysis of the images showed them to be of insufficient quality for an accurate magnitude and colour catalogue to be made.

Consequently, another 47 Tuc CMD dataset was used for this work. These
data were taken by Ken Freeman and Michelle Doherty (Freeman, K.C., private communication) with the 40" and WFI, hence providing the same field of view (although a slightly different pointing) to the main time-series dataset. These CMD data were then reduced as per the procedure described earlier and then the images in each filter were combined to produce a master V and master I image. By running DAOPHOT within IRAF on these two master images, the uncalibrated PSF magnitudes were obtained for 43,067 stars, which were then matched between the two colours. The output was a list of these stars with their positions on their respective CCD, their uncalibrated V magnitude and their uncalibrated V-I colour.

The raw output of DAOPHOT produced magnitude differences from one star to the next relative to an arbitrary reference magnitude. The zero-point of this output is unknown. Therefore the output catalogue needed to be calibrated in order to measure stars at their true magnitudes. This is usually performed by measuring a set of standard field stars, that is stars with accurately-known true magnitudes. In this case, using standard stars was not an option, as the data for these were not available. This was, however, not an unresolvable problem, as 47 Tuc is a well observed object, and many calibrated CMDs exist in the literature for this cluster (Hesser et al. 1987; Kaluzny et al. 1998). The procedure used to calibrate the CMD shall now be described, along with a presentation of the calibrated CMD dataset used throughout the rest of the thesis.

2.6.1 Calibration

The CMD dataset was calibrated initially by comparing the raw magnitude data to the previously published data of Kaluzny et al. (1998), who provided a wide-field photometric dataset of 47 Tuc online\(^4\). The authors warn in their paper of systematic errors caused by non-linearity of the OGLE CCD chip. For faint stars these errors are likely to be more significant. However, the stars needed for the calibration lie at the cluster turnoff, red giant branch and horizontal branch, and are of bright magnitudes; hence the dataset is quite suitable for this analysis. To calibrate the data, the V-magnitude and V-I colour distributions of the two datasets were overlaid and the uncalibrated data shifted until they best matched the Kaluzny data. Fig. 2.8 shows the distribution of V magnitudes between the two datasets, after the final calibration of the 40" CMD data in V. Using this simple method, the photometry in V is accurate to \(\leq 0.03\) mag when compared to previously published catalogues for the cluster, measured as the mean deviation of the two distributions.

As seen in Fig. 2.8, the red histogram was shifted by small increments (0.01 mag) until the two distributions best matched. It can be seen that the 40" data

\(^4\)http://cdweb.u-strasbg.fr/Abstract.html
Figure 2.8: The CMD calibration in V. The blue histogram is the distribution of calibrated 47 Tuc V magnitude data taken from Kaluzny et al. (1998), and the red histogram is the V distribution from the CMD dataset used in this thesis, shifted to match. This simple calibration process permits a dataset with a calibration accuracy of \( \leq 0.03 \) mag in V. The apparent spike in the red distribution seen at \( V=19.5 \) is caused by the superposition of 47 Tuc main sequence stars with Small Magellanic Cloud red giant branch stars at that magnitude.

does not reach as bright an upper limit as the Kaluzny et al. (1998) data, with fewer stars at \( V=14 \) at the location of the cluster Horizontal Branch. However, the two distributions match very well, accounting for the differences caused by the differing fields of view between the two datasets.

Similarly, Fig. 2.9 shows the distribution between the V-I colour of the Kaluzny et al. (1998) and 40\text{''} datasets. This is needed to calibrate the stars in I, and to obtain an accurate calibration in both I and V-I. The two distributions in V-I are very similarly matched, after the red distribution has been shifted until it best matches the blue. An apparent excess of stars is seen on the 40\text{''} distribution at V-I~0.8. This is attributed to the significantly wider field of view of the 40\text{''}, which samples much more of the cluster outer regions, and thus more cluster main sequence stars, leading to a pile-up at this colour typical of the cluster main sequence. From this distribution, the calibration in V-I for the 40\text{''} dataset is accurate to \( \leq 0.04 \) magnitude, again measured as the mean deviation of the two distributions. These errors are the nonrandom errors in the zero-point determination, and incorporate the errors in the OGLE calibration.

After performing this distribution matching process, Fig. 2.10 shows the results
Figure 2.9: The CMD calibration in V-I colour. Again, the blue histogram shows the calibrated V-I data from Kaluzny et al. (1998), and the red histogram is the 40" data shifted to match. These distributions are very similar, with the excess stars seen at V-I~0.8 being due to the 40" data having a substantially larger field of view, and thus more main sequence stars present in the outer regions of the cluster. Note the excellent correlation in the spike of the distribution.

of a further test of the CMD calibration. This figure shows the newly calibrated 40" data (black points) superimposed on a calibrated Stetson 47 Tuc dataset taken from the Canadian Astronomy Data Center\(^5\) (green triangles). While this Stetson data are mainly sensitive to the bright end of the scale, there is considerable overlap at the cluster turnoff. This is further evidence that the calibration of the CMD dataset is of suitable quality to place any transiting system or variable star on the standard V and I magnitude scale.

Fig. 2.12 shows the total calibrated CMD for 47 Tucanae, as used in subsequent chapters of this thesis. The total database contains 43,067 stars in the field, with V and V-I photometry to a photometric precision of 0.03 and 0.04 mag, respectively. The output DAOPHOT uncertainties are plotted as errorbars as a function of V. The spread of magnitudes in this dataset is clear, as is the cluster main sequence turnoff. The cluster horizontal branch is undersampled, as at V=14, it is expected that these stars will be saturated in the time-series dataset. As well as the obvious cluster stars, a further sequence can be seen intersecting the cluster main sequence at V=19.5, and running towards the red at brighter magnitudes. This is contamination from the background Small Magellanic Cloud.

\(^5\)http://cadcwww.hia.nrc.ca/cadcbin/wdb/astrocat/stetson/query/
Figure 2.10: To highlight the effectiveness of this simple calibration of the CMD dataset, the black points indicate the data for the brightest part of the CMD, and the green triangles indicate a further dataset used for conclusive comparison (Stetson). The degree to which both datasets overlap is easily seen, providing further proof that the calibrated CMD is consistent with previously published catalogues.

(SMC), and is the SMC red giant branch. The actual location of the intersection between these two sequences is seen as a clump of stars at $V=19.5$, $V-I=0.9$.

Fig. 2.11 shows the magnitude depth of our CMD photometry for all eight WFI CCDs, with recovered stars counted in 0.25 magnitude bins. Our photometry is limited to the extreme range $13.0 \leq V \leq 22.0$. It is clear that different WFI chips have somewhat different sensitivities. The apparent peaks in the stellar distributions are due to incomplete sampling at faint magnitudes. The gradual spatial decrease in star numbers towards CCD8 can be clearly seen, indicative of increasing distance from the Small Magellanic Cloud (SMC), and hence decreasing background contamination.

The number of stars in this dataset on the cluster main sequence is large, especially at the location just underneath the cluster turnoff. This is the location of the main transit search, and these data provide a high probability of obtaining $V$ and $V-I$ information for any transiting system identified. The magnitude limits
Figure 2.11: The magnitude limits of the CMD dataset, with the number of stars in each CCD plotted against their V magnitudes. It is clear that CCDs 1 and 8 indicate a fall-off in star numbers, due to the increased distance from the Small Magellanic Cloud. These limits give indications as to the range of magnitudes available for variable star identification.

of the data are apparent, and in summary this CMD dataset covers all stellar members of 47 Tucanae from low-mass, low-luminosity main sequence stars, to a few of the brightest red giant branch stars at the cluster horizontal branch and beyond. These data are of vital importance when discussing the expected depths and durations of the Hot Jupiter transits to be identified in the time-series dataset.

2.6.2 Astrometry

Astrometry was obtained for all 143,814 stars detected in our best seeing image, across the eight CCDs of the 47 Tuc field allowing a determination of the position of all stars in both our photometric and lightcurve database. We used a program to search the USNO CCD Astrograph Catalogue (UCAC1) (Zacharias
Figure 2.12: The calibrated CMD dataset for 47 Tucanae, as used in the rest of the thesis. These data are calibrated to an accuracy of 0.03 and 0.04 mag in V and V-I, respectively, and used to place any identified variable stars and transiting systems on the standard magnitude scale. The output rms uncertainty of the DAOPHOT photometry is plotted as errorbars. The cluster main sequence turnoff is clearly evident, as is a highly significant number of stars on the main sequence itself. As well as cluster stars, a large number of background Small Magellanic Cloud stars are seen as the second red giant branch intersecting with the cluster main sequence.

et al. 2000) for astrometric standard stars within the field (B.P.Schmidt, 2003 private communication). Several hundred such stars were cross-identified, allowing an accurate astrometric solution for our own stellar catalogue. The rms residual uncertainty of the astrometry was measured to be $\sim$0.15 arcsecs. Our astrometry is presented in Fig. 2.13, with the arrow indicating the direction to the core of the Small Magellanic Cloud. The cluster core and our extensive coverage of the cluster is apparent. With this astrometric solution, and the CMD presented earlier, it is now possible to obtain V magnitude, V-I colour and positional information for any star seen to undergo transits as well as for any identified variable stars.
Figure 2.13: The derived astrometry for the 47 Tuc field. Every point defines a star for which we have an astrometric solution with a measured rms uncertainty of 0.15 arcseconds. The degree of crowding towards the cluster core and the saturated inner core are apparent. With this dataset and the CMD data presented in Fig. 2.12, accurate knowledge of the V magnitude, V-I colour and astrometry can be determined for any transiting system or variable star observed. The arrow indicates the direction to the core of the Small Magellanic Cloud.

2.7 Chapter Summary

This chapter has described the observational strategies and data reduction techniques used to produce a dataset for 47 Tuc which has the best possible chance of allowing a detection of a Hot Jupiter transit. The choice of telescope and detector was made with this main aim in mind. The choice of filter and the determination of the optimum exposure time to maximise the chance of a detection was discussed. As appeared on the original telescope proposal, an estimate was made of the theoretical number of Hot Jupiters that should be contained in an
average dataset was made, and the resultant statistical significance estimated. These calculations show that given average conditions, the dataset should be able to provide a meaningful result.

The observing run was described along with the final dataset properties. Image reduction was undertaken on the images with standard reduction techniques. The cluster Colour Magnitude Diagram (CMD) was described, which is an important diagnostic tool for categorising the detected variable stars and transiting systems. The details of the CMD calibration and its photometric depth and accuracy were displayed. Finally, details of the astrometric solution were described, which, when combined with the CMD dataset, provides important information needed for follow-up observations.

The next step in the process is to produce the actual time-series from these reduced images. The way in which this was done is described in the next chapter, along with a description of the algorithm developed and used to search this large dataset for variable stars. Subsequent chapters deal with the variable star catalogue itself and details of the main planetary transit search.