Dissertation
submitted to the
Combined Faculties of the Natural Sciences and
Mathematics
of the Ruperto-Carola-University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Put forward by

Master Phys. Maximiliano Leonardo Moyano D'Angelo
born in: Santiago (Chile)

Oral examination: January 19th, 2011

A search for transiting extrasolar planets with the LAIWO instrument

Referees: Prof. Dr. Thomas Henning
Prof. Dr. Joachim Wambsganss

Zusammenfassung

In dieser Arbeit untersuchen wir die notwendigen Routinen, um extrasolare Planeten nach der Transit-Methode zu suchen. Wir wenden diese Methoden in einem der Himmelsfelder des LAIWO-Projekts im Cygnus-Lyra-Feld ("Laiwo VI") an und beschreiben, welche Probleme systematische Effekte für die präzise relative Photometrie auf dem Millimagnituden-Niveau ($\sim 3mmag$) darstellen. Ferner beschreiben wir Wege dieses korrelierte Rauschen zu quantifizieren und zu minimieren. Wir vergleichen die Stärken und Schwächen der zwei Transit-Detektions-Algorithmen "Box fitting algorithm" (BLS) und TRUFAS, indem wir Archiv-Daten des OGLE-Projekts und Simulationen von Daten aus dem ersten Jahr der Pan-Planets-Durchmusterung verwenden. Diese beiden Datensätze sind hinsichtlich Hauptspiegeldurchmesser und Gesichtsfeld der verwendeten Instrument-Teleskop-Kombination vergleich zum LAIWO-Survey. Wir fanden, dass der BLS-Algorithmus hauptsächlich durch die Transittiefe und das korreliertes Rauschen (das so genannte "Red Noise") beschränkt ist. Die TRUFAS-Detektionseffizienz korreliert mit der Anzahl der Messpunkte im Transit und der Anzahl der beobachteten Transits; sie ist kleiner als $\sim 50\%$ für die oben beschriebenen bodengebundenen Beobachtungen. Schließlich erzeugen wir Lichtkurven aus den LAIWO-Daten die zur Entdeckung von Planeten um Sterne heller als R = 16.5 geeignet sind. Darin fanden wir 31 bedeckungsveränderliche Doppelsterne und 18 Lichtkurven, die mit Planetentransits vereinbar sind. Drei dieser bedeckungsveränderlichen Doppelsterne und acht der Planetentransits wurden in der unabhängigen Durchmusterung von KEPLER gefunden. Unter den zehn von KEPLER nicht gefundenen möglichen Planetendurchgängen sind drei erfolgversprechende Kandidaten, die nun in einer Nachfolgestudie gründlicher untersucht werden müssen, um zu bestätigen, dass es sich bei den Transits tatächlich um Planeten handelt.

Abstract

In this thesis we study the necessary methods to perform a transit search for extrasolar planets. We apply these methods to search for planets in one of the fields of the LAIWO project: the Cygnus-Lyra field ("Laiwo VI"). We describe the problems that systematic effects can introduce for precise relative photometry at the millimagnitude level ($\sim 3mmag$). Ways to minimize and quantify this correlated noise are also described. We test the weaknesses and strenghts of two transit detection algorithms (TDA) namely the Box fitting algorithm (BLS) and the TRUFAS algorithm using archive data from the OGLE project and simulations of the first year of the Pan-Planets survey. These projects are similar in terms of telescope size and field of view to the LAIWO survey. We have found that the main limitations of the BLS algorithm are the transit depth and correlated noise ("Red Noise"). The TRUFAS detection efficiency correlates with the number of points in transit and the number of transits present in the light curve, and, its detection efficiency is low (less than $\sim 50\%$) for these type of ground-based observations. Finally, we create from the LAIWO data light curves which are suitable to detect planets among the stars brighter than R = 16.5 mag. We have found 31 eclipsing binaries and 18 light curves that have transits consistent with a planet. Of these detections, 3 eclipsing binaries and 8 planet candidates were independently found by the KEPLER survey. Of the 10 newly discovered transiting planets, 3 are promising to justify follow-up confirmation studies, which are always necessary to probe the planetary nature of a transiting companion.

Para mi gran padre Ricardo.

Contents

List of Figures

LIST OF FIGURES

List of Tables

1

Introduction

1.1 Introduction

The search for life on other planets has been a quest that has motivated human beings since the genesis of mankind. Are we alone in the universe?, Is there any form of life elsewhere? are questions which have intrigued scientists for decades. The first discovered extrasolar planets opened an encouraging opportunity to finally pursue the answer to some of these fundamental questions.

Life as we understand it is mainly possible because of liquid water, like the case of the earth (Chyba & Hand, 2005). Planets that are at a distance of their host star where liquid water can exist on the planet's surface can presumably harbor life. This ring around the star is called the habitable zone (HZ). To search for such earth equivalents is one of the main motivation of extrasolar planet searches.

Astronomers have conceived cunning techniques to detect extrasolar planets. In most cases, they rely on the influence of the planet on its parent star. These techniques have allowed to discover almost 500 extrasolar planets until now[1]. Briefly, some ways to detect extrasolar planets are:

- *Radial velocity measurements* : due to the mutual gravitational interaction between the planet and the star, the stellar radial velocity will change due to the motion of the star around the system's center of gravity. Variations in the stellar spectral lines will show a small Doppler shift (the star will appear to move towards or away from the observer). Measuring these radial velocity changes allows to confirm the presence of planets (Mayor

[1]For the most recent number count, see exoplanet.eu

& Queloz, 1995). Radial velocity measurements permit to estimate the orbital period and the eccentricity, but only a minimum mass of the planet ($m \times \sin i$) due to the unknown inclination of the system.

- *Astrometry* : A star with a planetary companion orbits around the system's center of mass. If the star's position is measured very precisely over time, it is possible to model that orbit and discover the perturber planet.

- *Microlensing* : when a massive foreground object passes very close to the line of sight of a background source star, it is possible to observe the deflection and magnification of the emission of the background source due to the gravitational potential of the foreground object (Einstein, 1915). If the lens foreground object is a star with a planet, the presence of the planet may be detectable due to a brief disturbance in the lensing light curve (Bond et al., 2004; Mao & Paczynski, 1991). The microlensing technique allows to constrain the ratio of planet mass to stellar mass.

- *Pulsar timing* : pulsars emit radio waves whose timing can be measured very precisely. Due to the extremely regular rotation of a pulsar, small changes in the timing of the radio pulses can be used to define the pulsar's motion. If there is a planet orbiting the pulsar, it will interact gravitationally with the pulsar, and, therefore, influence its orbit (Wolszczan & Frail, 1992). This method is so sensitive that is capable of detecting planets as small as a tenth the mass of Earth.

- *Direct imaging* : A star's light reflected off a planet can be imaged directly. To probe that the object is associated with the parent star, proper motion measurements are necessary too (Kalas et al., 2008).

The *transit method* is a very promising method to detect extrasolar planets and has a certain advantages compared to the other techniques. When a system is observed nearly edge-on, it is possible to observe a periodic dim of the stellar light as the planet orbits. A transiting system allows to study several important properties of the system such as planetary radius, mass, and therefore also its density which is not possible to determine by any other methods. In combination with the radial velocity technique, the unknown inclination of the system can be resolved and the true planetary mass (not a minimum mass) can be estimated (Seager & Mallén-Ornelas, 2003). Moreover, transiting systems are the only ones that currently allow direct studies of planetary atmospheres. An earth-like planet transiting their parent star can

be identified unambiguously as terrestial. Thus it is conceivable to study the atmosphere of a transiting earth-like extrasolar planet. In an M dwarf star, the HZ is much closer to the star than in the solar system. Consequently, to observe transits of terrestial planets in the HZ is possible (Nutzman & Charbonneau, 2008). In addition, an M dwarf star has numerous advantages to detect transits: they would be more likely, more frequent, and present a much larger signal (see chapter 2).

In conclusion, it is now possible to discover earth-like planets in the HZ of their stars using the transit method. Such planets, could, in principle, harbor life. To search for earth-like equivalents using the transit method is, therefore, extremely interesting.

1.2 Outline of the thesis

This thesis is structured as follows:

Chapter 2 gives a general overview of the transit method. It describes all the important properties we can learn from a system with a transiting planet. In addition, it gives a glimpse of some promising follow-up studies that are possible for transiting systems, which could allow in the near future the first studies of earth-like planet atmospheres outside our solar system.

Chapter 3 describes the problem of correlated noise in transit surveys. It summarizes the main obstacles that systematics present to the detection of transiting planets. The way to minimize and quantify the systematics is discussed. It is crucial to remove the systematics to detect unambiguously small brightness decreases due to transiting planets.

Chapter 4 summarizes some transit algorithms designed to detect shallower transits produced by planets. Some improvements of these techniques are also discussed.

Chapters 5 and 6 describe the application of the transit detection algorithms (TDA) to real and simulated data, the OGLE and Pan-Planets data sets respectively. They are ideal data sets to test the TDA and evaluate their efficiency. Both data sets are very similar to the LAIWO data.

Chapter 7 summarizes our results for the Cygnus-Lyra field, one of the monitored fields of the LAIWO project. All the steps required to build high quality light curves are described. The data calibration process, high quality photometry, and removal of systematic effects are detailed. Finally, the search for planets and selection of promising candidates is presented.

Chapter 8 recaps our results, conclusions, and the future work after this dissertation.

2

The Transit Method

2.1 Overview

When a planet crosses ("transits") the line of sight between the observer and the planet's host star it is possible to observe a periodic small drop in brightness as the planet orbits (see figure 2.1). The amount the star dims depends on the relative sizes of the star and the planet, it is usually below $\sim 1.5\%$ ($1\% \sim 0.01$ mag). For example, a Jupiter-like planet orbiting a Sun-like star decreases the stellar light in $\sim 1\%$.

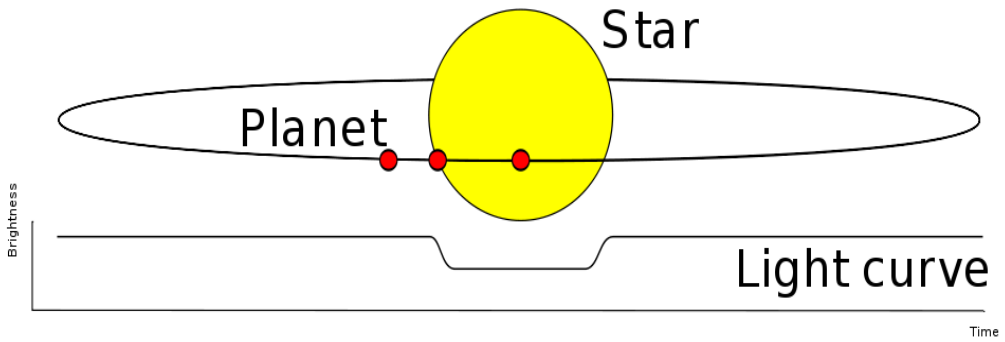

Figure 2.1: Schematic of a planetary transit. When the planet transits, a small dip in the stellar light is observed.

Assuming random orientations of a planet's orbit with respect to the line of sight, the geometric probability of a planetary system to show a transit is given by $\mathbb{P}_{tr} = R/a$, where R is the stellar radius and a is the orbital semi-major axis (Sackett, 1999). For a Jupiter-like planet orbiting a Sun-like star this probability is 1/1100. If we assume that 1% of the stars

have a Jupiter, it is possible to observe one transit every 12 years (Jupiter's period) if 100,000 stars are monitored, therefore, it was not encouraging to use the transit method to search for planets. The discovery of the "Hot Jupiter" class of planets by radial-velocity measurements in 1995 (Mayor & Queloz, 1995) improved dramatically the prospects of detecting planets with the transit method. For such systems, the transit probability is $\sim 10\%$, which implies that the detection of a Jupiter-like planet is possible if few thousands stars are monitored, within easy access of modern wide-field CCDs. In addition, the transits in such systems are more frequent ($\sim$ few days), thus it is easier to discover them. Several transit surveys started to operate motivated by the newly discovered class of Jupiters. Surveys such as OGLE (Udalski et al., 2002a), TRES (Brown & Charbonneau, 2000), HAT (Bakos et al., 2004), Super-WASP (Pollacco et al., 2006), and XO (McCullough et al., 2005) started to monitor the sky to search for transiting planets. The transit method had its first success in 1999, when the transit of the radial-velocity planet HD 209458b was observed (see figure 2.2, Charbonneau et al. 2000). Later in 2003, the planetary nature of the candidate OGLE-TR-56 (Konacki et al., 2003) was confirmed, thus it became the first planet discovered by the transit method. Currently, space-based missions like COROT (Baglin et al., 2002) and KEPLER (Borucki et al., 2010) are delivering several new interesting planets.

The transit method is bias towards detecting short period planets because the probability is higher. This can be clearly seen in figure 2.3, which shows that most of the detected transiting planets have short periods (less than 10 days). Surprisingly, it is also possible to observe transits caused by planets with long periods, like the case of HD 80606 (out of scale in figure 2.3), whose period is 111 days (Moutou et al., 2009).

The transit method is promising because the light curve allows us to access important properties of the system, such as the planetary radius (if the mass and stellar radius are known), and orbital inclination (Seager & Mallén-Ornelas, 2003). In combination with radial velocity measurements (a necessary step to confirm the planetary nature of the companion), transit measurements give us a direct estimation of the mass of the transiting body. In addition, a transiting planet offers the system configuration for the strongest radial velocity signal to be measured.

Currently, more than 100 transiting planets have been discovered. Figure 2.4 shows the current mass-radius (M-R) diagram for the known transiting extrasolar planets [1], one of the most important results of the transit method. It is only for transiting planets that we have

[1] See exoplanet.eu

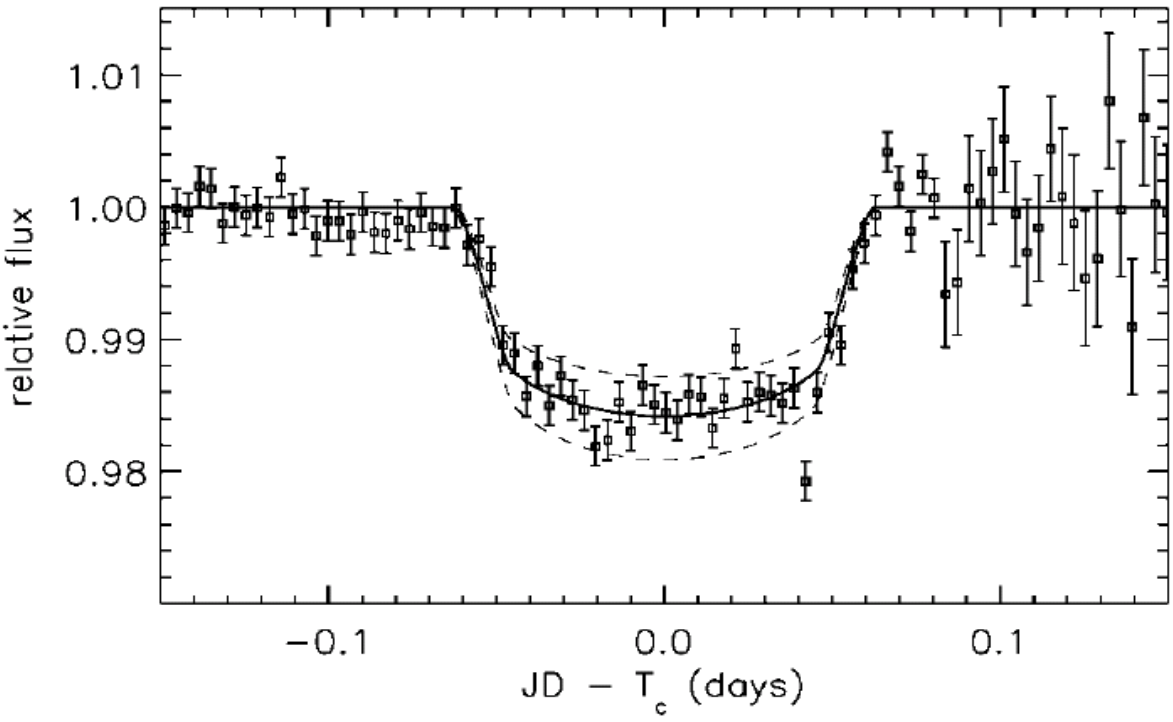

Figure 2.2: Light curve of HD 209458, the first observed planetary transit (Charbonneau et al., 2000).

direct estimates of their masses and radii. This diagram is crucial to put constraints on models of the internal structure of giant planets (Guillot, 2005).

2.2 Transit parameters

Figure 2.5 shows the basic observational parameters that define a transit: the period P, the transit duration Δt, the transit depth δ, and the transit epoch t_0.

The transit depth is related to the ratio of planetary to stellar radii (if we neglect limb-darkening, Seager & Mallén-Ornelas 2003):

$$\delta = \left(\frac{R_p}{R_\star}\right)^2 \tag{2.1}$$

For example, a Jupiter-like planet orbiting a Sun-like star produces a transit with a depth of $\sim 1\%$. The transit duration (for a circular orbit) is related to the orbital parameters (semimajor

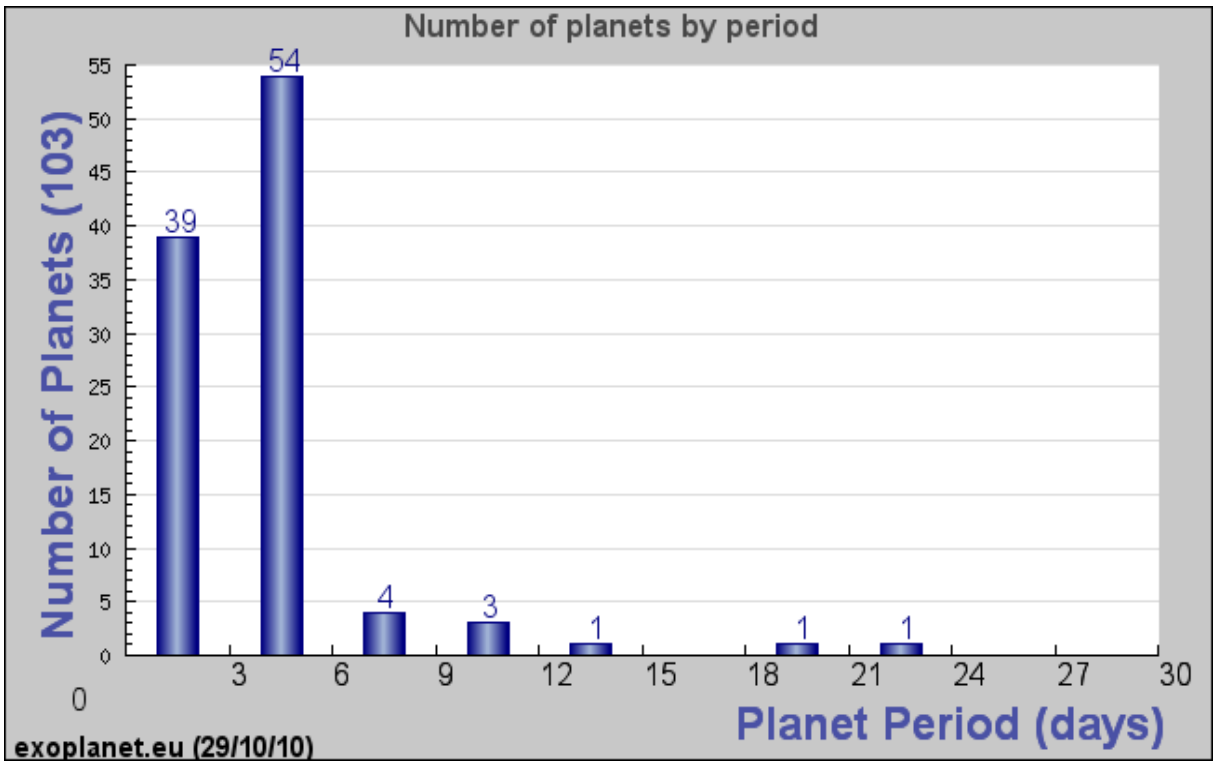

Figure 2.3: Histogram of the discovered transiting planets (see exoplanet.eu).

axis a, period P, and inclination of the system i) and to the stellar and planetary radius (Seager & Mallén-Ornelas, 2003):

$$\Delta t \simeq \frac{PR_\star}{\pi a}\sqrt{(1+\frac{R_p}{R_\star})^2-(\frac{a}{R_\star}\cos i)^2} \tag{2.2}$$

It is also possible to describe the transiting system using the ingress time (defined as the time when the planet starts eclipsing till it is totally covering the star) and the total duration to obtain an estimation of the density of the planet (using the Kepler's third law and the stellar mass-radius relation). For more details see Seager & Mallén-Ornelas (2003).

2.3 False positives

Stellar configurations or random noise can mimic a planetary transit. They outnumber the real transits due to planets (Brown, 2003). Figure 2.6 summarizes the most common confusion cases. In brief, they are:

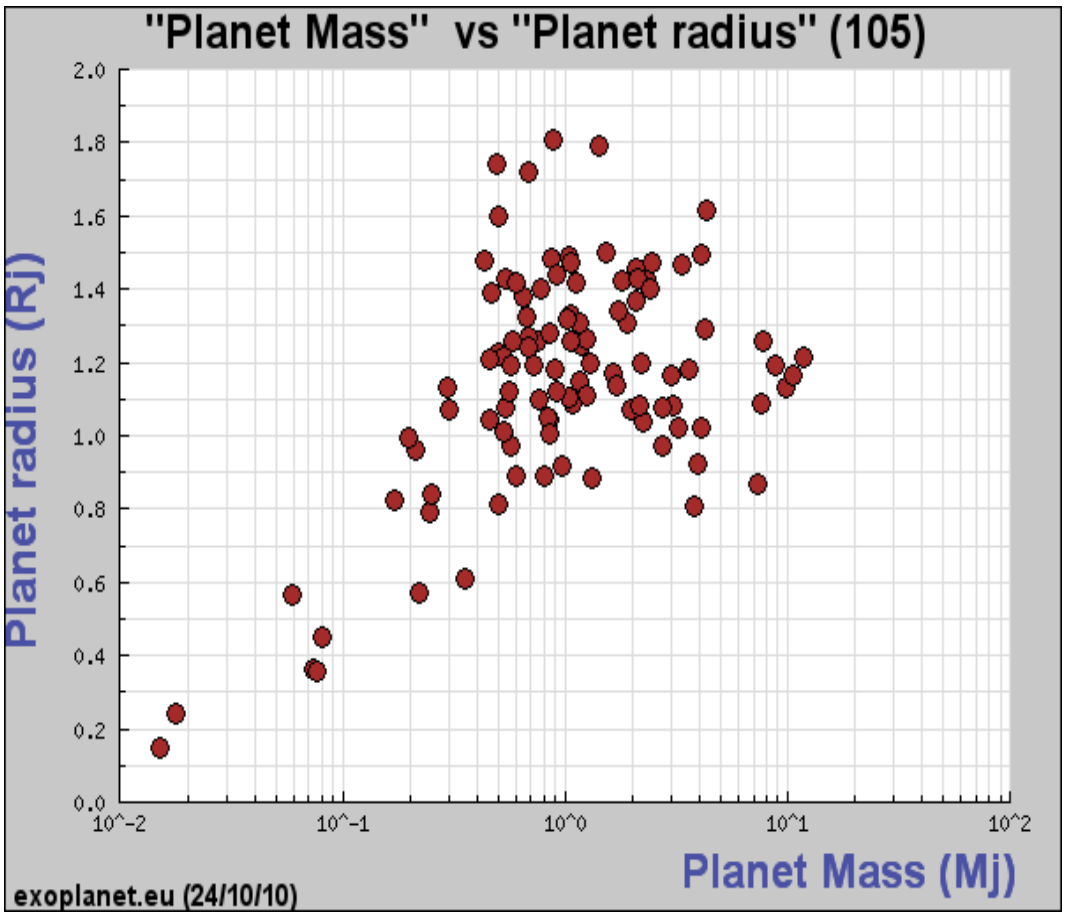

Figure 2.4: Mass-radius digram for the currently known transiting extrasolar planets (see exoplanet.eu).

- *Grazing binaries.* If binary stars orbit in an inclined angle, they could mimic a shallow transit-like dip in the light curve (*a)* in figure 2.6).

- *Small stellar companion.* A small M-dwarf or brown dwarf orbiting a larger star can produce a transit that resembles a planetary transit (*b)* in figure 2.6).

- *Eclipsing binaries with a third star.* A foreground star can dilute the light of a binary system and imitate a transit (*c)* in figure 2.6).

- *Statistical false positive.* Noise due to instrumental effects or observing conditions can also produce artifacts that can be confused with a planet-like transit (*d)* in figure 2.6).

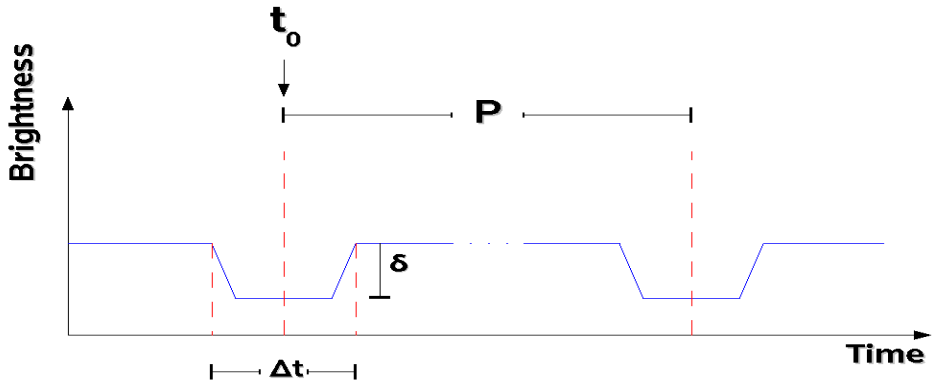

Figure 2.5: Transit light curve parameters. Here P is the period, Δt is the transit duration, δ is the transit depth, and t_0 is the epoch.

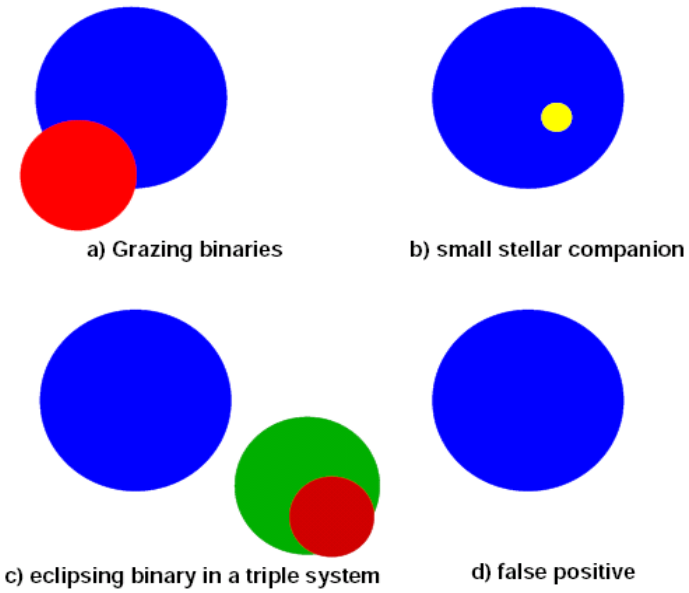

Figure 2.6: Schematic of common false positives (Moutou et al., 2006). a) grazing binaries, b) binaries with a small companion, c) binaries with a blended third star, d) statistical false positive.

There are ways to rule out some false positives using the light curve alone. Transits with V-shape (grazing binaries) or ellipsoidal variations are clear indicators of a massive companion (Sirko & Paczyński, 2003). Photometry at different passbands (Tingley, 2004) or a transit duration not consistent with planetary size (Tingley & Sackett, 2005) are also ways to select good transit candidates for radial velocity confirmation.

2.4 Follow-up Studies

In addition to the direct estimation of planetary radii and mass, and, therefore the density, a transiting planet awards the possibility of studying several other interesting features of the system. Currently, a planet transiting their star is the only system configuration that allows direct studies of the chemistries and dynamics of planetary atmospheres. We will briefly describe some of the possible follow-up studies that transiting systems afford.

2.4.1 Transmission Spectroscopy

During transit, some of the stellar light crosses the planetary atmosphere. Differences in spectra between in and out of transit phases reveal wavelength dependent features, that indicate the presence of atomic or molecular species in the planet's atmosphere. These features reveal different heights in the planetary atmosphere causing variations from the stellar continuum of $\sim 0.1\%$ (Seager & Sasselov, 2000). The first success of this technique was in 2002, when Charbonneau et al. (2002) discovered Sodium in the atmosphere of the transiting planet HD 209458 (Charbonneau et al., 2000).

Figure 2.7 shows the detection of Methane and Water in the atmosphere of the transiting planet HD 189733 (Bouchy et al., 2005b; Swain et al., 2008) using the transmission spectroscopy technique at infrared wavelengths.

2.4.2 Occultation Spectroscopy

When a transiting planet is eclipsed by the star, it is possible to observe a small dip in the star light ("secondary eclipse"). The missing flux corresponds to photons emitted by the planet blocked by the star. Measurements of the secondary eclipse depths at different wavelengths (usually at the infrared) are estimators of the planet-to-star flux ratios (Charbonneau et al., 2005; Deming et al., 2005). These technique allows the detection of thermal emission radiated by the planet and helps to constrain theoretical models of the planetary atmospheres.

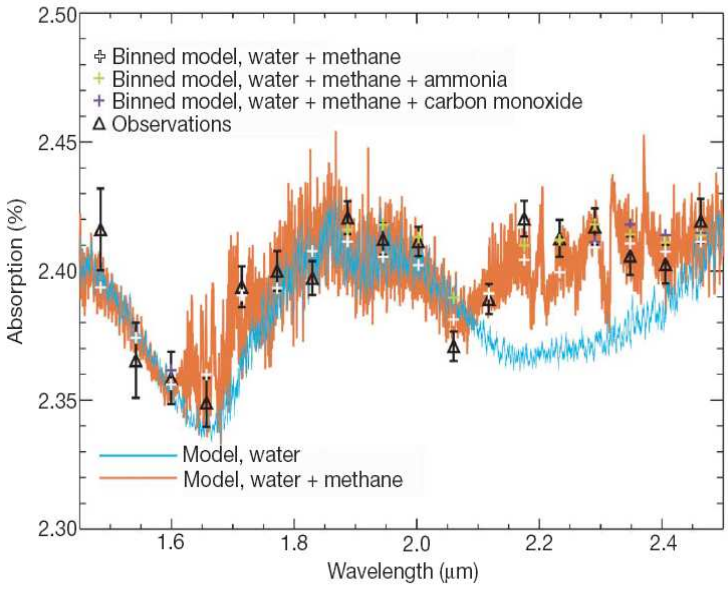

Figure 2.7: Tranmission spectrum of the transiting planet HD 189333 (Swain et al., 2008). The presence of Water ($\sim 1.9\mu m$) and Methane ($\sim 2.2\mu m$) is confirmed.

Figure 2.8 shows the wavelength dependence of the secondary eclipse depths of the planet HD 189733 (Bouchy et al., 2005b; Charbonneau et al., 2008). Figure 2.9 shows the estimated planet-to-star flux ratios with two models superimposed, one where the emission of the absorbed stellar flux is constrained to the day side (upper curve) and another where the energy is distributed over the entire planet (bottom curve). The dashed line is a Planck curve with a temperature of 1292 K, which is a poor fit, therefore, it probes that there are spectral variations due to the emission of the planet (mainly water and CO).

2.4.3 Spin-Orbit alignment

When the planet is transiting the disk of its host star, it occults part of the stellar rotating surface, which produces an alteration of the stellar line profiles. Therefore, the planet removes a stellar velocity component from the broadened stellar absorption lines as it orbits. If the planet blocks a blue part of the line, the line will appear to be redshifted and vice versa. This effect is known as Rossiter-McLaughlin effect (McLaughlin, 1924; Rossiter, 1924).

Observing this effect allows to measure the line-of-sight velocity of the stellar disk ($v \sin I_\star$), and the angle λ between the sky-projected angular momentum of the planetary orbit and the stellar spin (Winn et al., 2005).

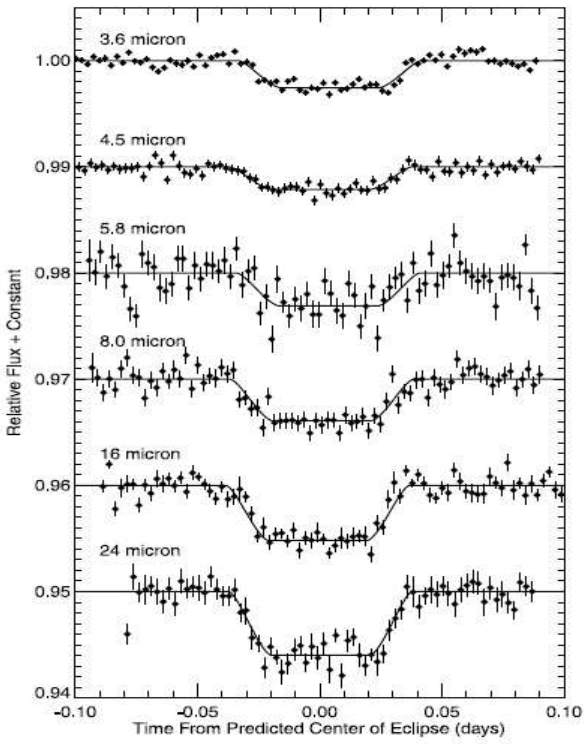

Figure 2.8: Photometry of HD 189733 at times when the planet is passing behind the star. The secondary eclipse observations are the *Spitzer Space telescope* infrared bands at 3.6, 4.5 , 5.8, 8.0, 16, and 24 μm using the IRAC instrument, MIPS instrument, and IRS instruments. The best-fit eclipse curves are overplotted (Charbonneau et al., 2008).

Figure 2.10 shows an example of the Rossiter-McLaughlin effect measured in the planet HD 209458 (Winn et al., 2005).

2.4.4 Transit timing variations

Measurements of the variations of the predicted transit times in already known transiting systems allow the detection of additional planets. These additional planets (not necessarily transiting) interact gravitationally with the transiting planet, producing a time-variation depending on the mass of the perturber. In some cases, terrestrial-mass planets can produce a measurable effect (Holman & Murray, 2005). The time variations permit the determination of the orbital period and mass of the perturber from transit observations alone.

During the orbits, the planets exchange energy and angular momentum due to their gravitational interaction, which produces oscillations of the semimajor axes and eccentricities of the

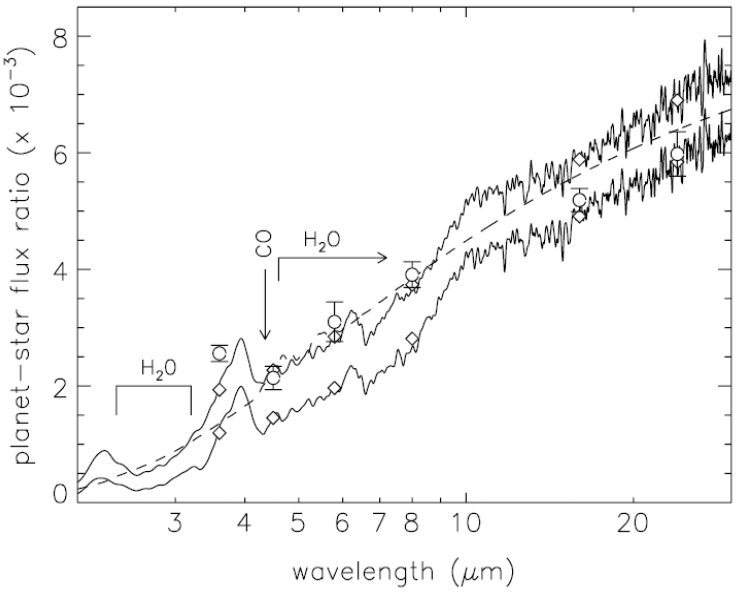

Figure 2.9: Comparison of the *Spitzer* infrared planet-to-star flux ratio measurements and models. Solid curves show the two models, one with the energy constrained to the day side of planet (upper curve) and the other with uniform energy distribution (bottom curve). The dashed line shows the Planck spectrum with a temperature of 1292 K (Charbonneau et al., 2008).

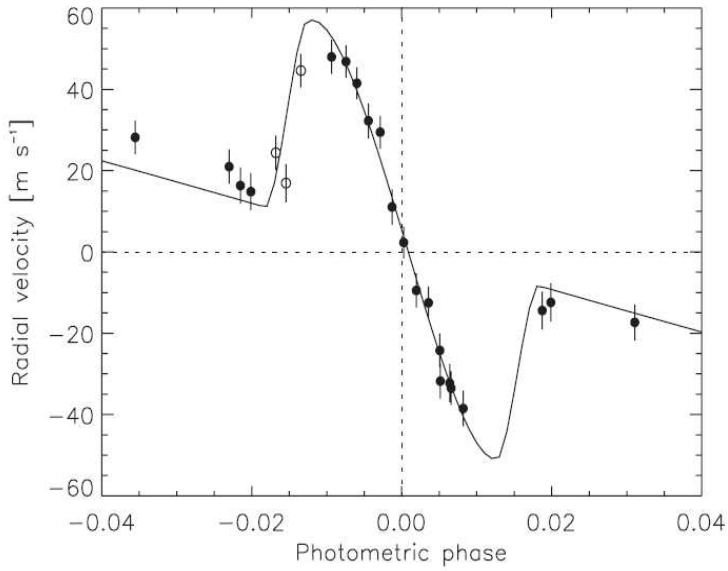

Figure 2.10: Rossiter-McLaughling effect observed on the planet HD 209458 (Winn et al., 2005).

planets, which cause variations in the time intervals between successive transits.

In cases in which two or more planets are transiting, the masses and radii of each planet can be estimated. Many stars are too faint for radial velocity follow-up, thus the transit-timing technique may be the only way to estimate the masses and densities of the planets for such systems (Holman & Murray, 2005).

Figure 2.11 shows the transit timing variations measured in the Kepler-9 system (Holman et al., 2010). The first system discovered with multiple transiting planets (Kepler 9b, 9c). The planets are most likely in 2:1 mean motion resonance (MMR), which implies that the period derivatives have opposite sign (effect observed in figure 2.11). The interdependent timing variations reveal that the two bodies are gravitationally interacting and, therefore, must be orbiting the same star.

2.5 Summary

A transiting system gives access to several important properties of the system, such as planetary mass and radii, which allow to constrain models of the internal structure of extrasolar planets.

Several stellar impostors can mimic a transiting system, thus the radial velocity confirmation is a must to confirm the planetary nature of the object.

Transiting systems give the opportunity to study the atmospheres and thermal emission of the planets. In addition, studying perturbations in already known systems allows to discover new perturber planets.

Space-based missions like KEPLER can discover the first earth-like transiting planets, thus, to study the atmospheres of such planets could allow the first attempts to search for life in other planets.

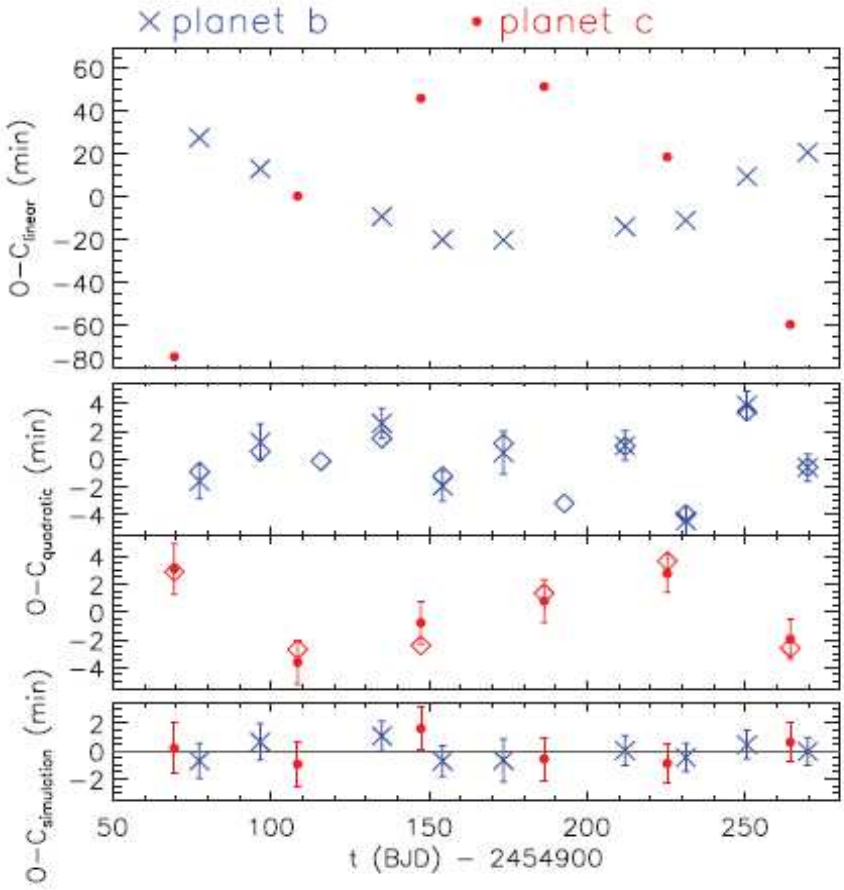

Figure 2.11: Transit timing variations in the KEPLER-9 system (Holman et al., 2010). Kepler *9b* and *9c* are indicated in blue and red symbols respectively. The observed (O) minus the calculated (C) values of the transit times for linear (top panel) and for quadratic (middle panel) ephemerides are shown. Diamonds indicate the dynamical model in which the planets fully interact. Bottom panel shows the comparison between radial velocity measurements and the dynamical model (Holman et al., 2010).

3

Systematic Effects - Red Noise

Systematic effects are always present at a certain level in photometric surveys. They can be intrinsic to the data, due to uncorrected instrumental effects/observing conditions, or they can originate from data reduction. In transit surveys searching for planets, these effects are significant and, in fact, they can completely limit the survey detection capability (the usual desired photometry precision must be better than 1.5%). This correlated noise is called "Red Noise" (Pont et al., 2006). These trends are in general related to changing airmass, atmospheric conditions, telescope tracking, flat-field errors, or more likely a combination of these factors. Assuming only white noise leads to an overestimation of the real signal-to-noise ratio of the transit (Pont et al., 2006) which can, in principle, cause the detection of false candidates. In fact, under certain conditions the red noise can completely dominate the uncertainties. In summary, it is important to understand and minimize the systematics in a transit survey searching for planets.

3.1 Red Noise Removal

Since the discovery of the first transiting exoplanets more than a decade ago, several techniques to remove systematics have been proposed. It is known now that the correlation of photometric time series at the millimagnitude level cannot be neglected. In what follows, some methods to remove and quantify the systematic effects will be described.

3.1.1 Atmospheric refraction systematics

One of the first attempts to correct systematics present in transit light curves was presented in Kruszewski & Semeniuk (2003). They noticed that small systematics appeared when measuring magnitudes using fixed positions without taking into account the atmospheric refraction. In crowded fields the blending is not uncommon, thus to have two unresolved stars of different colors or a single star contaminated by neighbours is possible. The refraction is color dependent (Monet et al., 1992), therefore, the central position of this unresolved system changes and the resulting magnitude is affected. Using the OGLE-III set (Udalski et al., 2002a) they showed that this effect is not negligible (see Fig. 1 of Kruszewski & Semeniuk 2003) and they proposed a polynomial dependence of magnitude on hour angle and time (or phase) to decrease the systematics due to refraction (see equations 2 and 3 of Kruszewski & Semeniuk 2003).

3.1.2 Sysrem

SYS-REM (systematics removal, Tamuz et al. 2005) was motivated initially by the color-dependent atmospheric extinction. The colors are not always known, which is sometimes the case of photometric surveys when only one filter is used and no color information is available.

Consider a set of N light curves, each of which consists of M measurements. Let $\{r_{ij}, i = 1, \ldots, N,\ j = 1, \ldots, M\}$ be the average-subtracted stellar magnitude of the star i in image j. Let $\{a_j, j = 1, \ldots, M\}$ be the airmass at which the j image was observed. We can now calculate the best linear fit for the residuals of star "i" and define the effective extinction coefficient $\{c_i, i = 1, \ldots, N\}$ as the quantity that minimizes the expression:

$$S_i{}^2 = \sum_j \frac{(r_{ij} - c_i a_j)^2}{\sigma_{ij}^2} \tag{3.1}$$

where σ_{ij} is the individual uncertainty of measurement of star i in the image j.

A simple differentiation of equation 3.1 gives an estimate of the extinction coefficient:

$$c_i = \frac{\sum_j (r_{ij} a_j / \sigma_{ij}^2)}{\sum_j (a_j^2 / \sigma_{ij}^2)} \tag{3.2}$$

The problem can now be turned around. We can determine the most suitable airmass of each image, using the known extinction coefficients of each star. Therefore, we want to minimize:

$$S_j{}^2 = \sum_i \frac{(r_{ij} - c_i a_j)^2}{\sigma_{ij}^2} \tag{3.3}$$

Given the previously calculated set of $\{c_i\}$, the value of the effective airmass is:

$$a_j^{(1)} = \frac{\sum_i (r_{ij} c_i / \sigma_{ij}^2)}{\sum_i (c_i^2 / \sigma_{ij}^2)} \tag{3.4}$$

It is possible now to recalculate the coefficients $c_i^{(1)}$ and continue iteratively. This is in essence a repetitive process that searches for the best two sets $\{\bar{c_i}\}$ and $\{\bar{a_j}\}$ that minimize the global expression:

$$S^2 = \sum_{i,j} \frac{(r_{ij} - c_i a_j)^2}{\sigma_{ij}^2} \tag{3.5}$$

Simulations of Tamuz et al. (2005) showed that this iterative process converged to the same $\{\bar{c_i}\}$ and $\{\bar{a_j}\}$, no matter what initial values were used. The values of the final set of parameters $\{\bar{c_i}\}$ and $\{\bar{a_j}\}$ are not necessarily related to the true airmass and extinction coefficients, they are simply the variables by which the global sum S^2 varies linearly most significantly. They could represent any systematic effect hidden on the data as long as the global minimum of S^2 is achieved.

It is natural now to define $r_{ij}^{(1)} = r_{ij} - \bar{c_i}\bar{a_j}$ and search another linear effect in the residuals $r_{ij}^{(1)}$. In summary, this process can be applied repeatedly, until it finds no significant linear effects in the residuals.

Formally, the process of identifying additional systematic effects can be repeatedly applied until there is no variation left in all light curves. A stopping criterion is needed to avoid removing real variability (transits, eclipsing binaries, etc.). A criteria is proposed in Mazeh et al. (2007), which is based on the fraction of light curves that show a scatter change below certain cut (denoted β_{min}). We have run SYSREM in LAIWO light curves (see chapter 7) and applied a slightly different stopping criterion, which is based on the mean scatter difference improvement of the light curves rather than a scatter cut (for more details see section 7.4).

3.1.3 Trend Filtering algorithm

The Trend Filtering Algorithm (TFA) (Kovács et al., 2005) uses the principle that systematic variations in a given light curve are shared by many light curves of other stars in the same data set. Thus, it is possible to identify objects in the field that suffer from the same type of systematics as the target, and correct the trends using some kind of filtering.

It is assumed that there are enough objects to represent all possible systematics. Once this subsample (also called a "template set") is selected, a filter function is built from these light

curves. This filter function is subtracted from the other, non-template light curves to remove the systematics.

The template set is built with objects uniformly distributed along the field. The idea is to ensure uniform sampling of the parameter space (position, color, etc.). It is then necessary to define a time-base where the filter function will be applied. This is usually not a problem because it is common to measure all stars over the same time-base. If measurements of some of the template objects are missing, they are filled with their averages. Finally, the template objects are zero-average using some outlier selection criterion (it is recommended to use $5 \times \sigma$ clipping).

The filter function is chosen as a linear combination of the template light curves. The objects of the template set are weighted based on their variance, since most of them are presumably non-variable. This linear combination is created separately for each target light curve from the template set. Thus if there are variables in it they will have a small effect, because their variation is most likely not correlated with systematics one is trying to filter.

3.2 Quantifying the Red Noise

Even after the red noise removal, small scale systematics remain in the light curves. It is possible to quantify this red noise level and calculate the signal-to-noise ratio (SNR) of the transit taking into account the systematics (Pont et al., 2006).

For a given light curve, the SNR of the transit including red noise can be calculated as:

$$S_r{}^2 = d^2 \frac{n^2}{\sum_{k=1}^{i=N_{tr}} n_k{}^2 \nu(n_k)} \tag{3.6}$$

Here d is the transit depth, n is the total number of points in the transit, N_{tr} is the number of transits, n_k is the number of points in the k-transit and $\nu(n_k)$ is the covariance function. It is assumed that the covariance between points of different transits will be much smaller than that for points of the same transit, because $\Delta t \ll P$ (where Δt is the transit duration and P is the period). The uncertainty on the depth of a single transit will be given by $\nu(n)^{1/2}$. The covariance function can also be expressed as two contributions, one purely white component, noted σ_w, and a purely red component, noted σ_r (where *'w'* and *'r'* stand for 'white' and 'red' respectively). For purely white noise, the variance is given by $\nu(n) = \sigma_w^2/n$, while for purely red noise it scales as $\nu(n) = \sigma_r^2$. In the general case, the covariance can be modeled as $\nu(n) = \sigma_w^2/n + \sigma_r^2$. It is easy to see from the previous expression that even with high number of points in transit the

signal-to-noise ratio is dominated by systematics if the red noise level high, i.e. adding more points to the transit doesn't give more information.

It is possible to estimate the covariance function $\nu(n)$ for a given transit using the light curve alone. This is done by relating the $\nu(n)$ to the variance of the average of "n" points in a time interval Δt outside the transit signal. The recipe is given in section 2.6 of Pont et al. (2006).

3.2.1 Red Noise in OGLE

Following the procedure of section 2.6 of Pont et al. (2006) we have estimated the covariance function $\nu(n) = {\sigma_w}^2/n + {\sigma_r}^2$ and calculate the values for σ_w and σ_r for every candidate of the OGLE survey (Udalski et al. 2002a, 2003, 2002b, 2004, 2002c, see chapter 5)

Figure 3.1 and 3.2 show examples of our fit for the candidates OGLE-TR-17 and OGLE-TR-7 (see Figure 5 of Pont et al. 2006). In both cases, there is a contribution due to correlated noise, the points don't follow a $\sigma/\sqrt{n}$ relation (dotted line). Clearly the red noise contribution is higher for the candidate OGLE-TR-7 than for the candidate OGLE-TR-17. This is explained because for faint stars the photon noise dominates. For bright stars the red noise is the dominant component. Figure 3.3 shows the fitted values σ_w and σ_r for every transit candidate of the OGLE survey (see Figure 8 of Pont et al. 2006). The amplitude of the white noise (blue circles) depends linearly on magnitude, because photon noise is dominating, except at the bright end. The red noise (red squares) has no clear dependence on magnitude. In fact, values of σ_r for different targets are really similar. Thus, it is a good approximation to consider that all the targets in a transit survey are affected by a similar red noise component. The average of σ_r for the OGLE candidates is 3.2 mmag. In conclusion, the red noise dominates in bright stars for the OGLE survey, but there are some exceptions that are clearly seen in Figure 3.3. Pont et al. (2006) found that a mean red noise component of $\sim$ 3mmag is typical value for other transit surveys as well.

3.2.2 Red Noise in Pan-Planets

We have estimated the red noise contribution on $\sim$ 100,000 Pan-Planets simulations (Koppenhoefer et al. (2009), see chapter 6). Table 3.1 summarizes the results. The global average of the red noise per transit of all simulated light curves is 2.3 [mmag], which is consistent with the level of red noise introduced in the simulations (see Koppenhoefer et al. 2009). We have also calculated the S/N taking into account the red noise (see equation 3.6) for all simulated light curves. These results are presented in section 6.4.4.

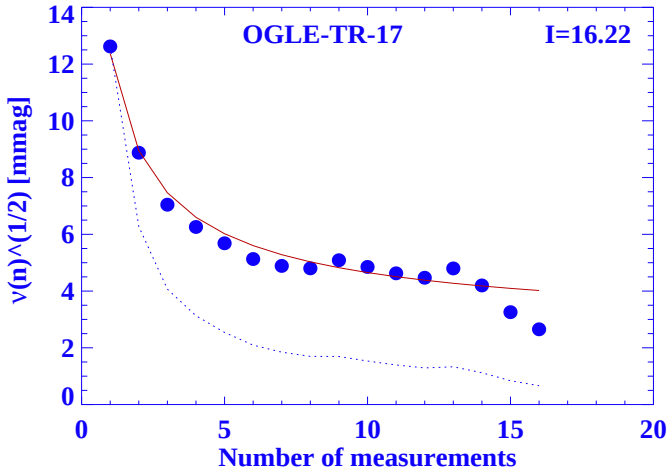

Figure 3.1: The $\nu(n)$ function for the candidate OGLE-TR-17. The $\nu(n)^{1/2}$ (standard deviation of n points) is plotted as a function of n. The red solid curve is a fit of $\nu(n) = {\sigma_w}^2/n + {\sigma_r}^2$. The dotted line shows a $\sigma/\sqrt{n}$ relation, as expected for white noise only. The mean magnitude is indicated in upper right corner.

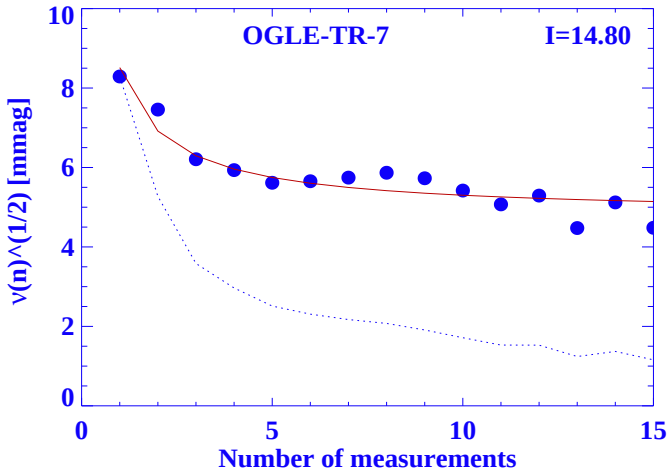

Figure 3.2: The $\nu(n)$ function for the candidate OGLE-TR-7. The $\nu(n)^{1/2}$ (standard deviation of n points) is plotted as a function of n. The red solid curve is a fit of $\nu(n) = {\sigma_w}^2/n + {\sigma_r}^2$. The dotted line shows a $\sigma/\sqrt{n}$ relation, as expected for white noise only. The mean magnitude is indicated in upper right corner.

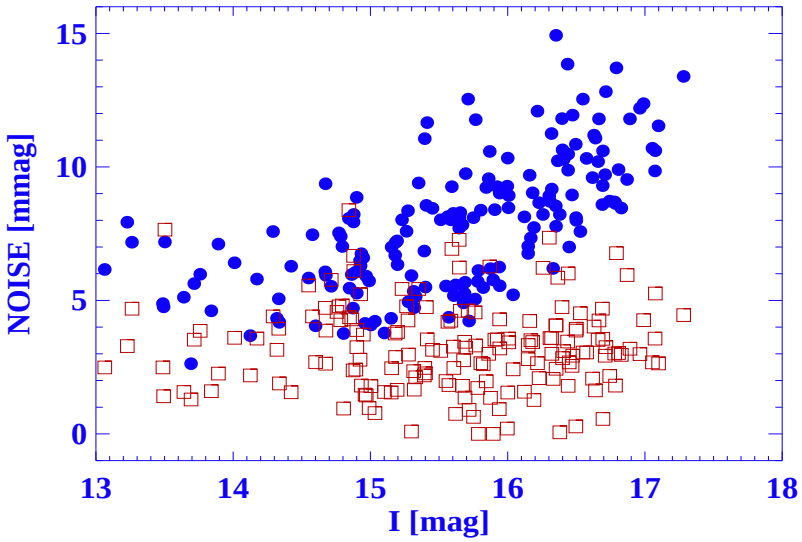

Figure 3.3: White and Red noise contributions for every OGLE candidate. Blue circles indicate the white component . Red squares indicate the red component. The white noise depends linearly on magnitude, while the red noise doesn't show any clear dependence on magnitude. The average red noise for the OGLE survey candidates is 3.2 mmag.

Table 3.1: Red noise in Pan-Planets

Simulation ID[a]	Red noise [mmag]
1	2.30
2	2.12
3	2.41
4	2.18
5	2.48
6	2.21

[a] For the details of the simulations see table 6.1.

3.3 Conclusions

To understand and remove the systematics in a transit survey is crucial to detect shallower transits due to planets.

The quantification of the red noise in the OGLE survey and Pan-Planets simulations show

that there is always residual red noise component at the millimagnitude level (average $\sim$ *few mmag*). The systematics are dominant for bright stars in these kind of ground-based surveys.

The red noise must be taken into account when evaluating the significance of a transit detection.

4

Transit Detection Algorithms

4.1 Introduction

Several transit detection algorithms have been proposed in the literature (see Moutou et al. 2005; Tingley 2003 for an overview). The first theoretical comparison was done by Tingley (2003); Tingley (2003). The main conclusion was that "no detector is clearly superior for all transit signal energies" but an optimized BLS algorithm (Kovács et al., 2002) still performs slightly better for shallower transits. The second comparison was done in the context of the next generation space-based transit surveys like CoRoT (Baglin, 2003). Five independent methods of analysis of a thousand synthetic light curves were presented in Moutou et al. (2005). This was the so called "CoRoT Blind Test 1". The light curves were built by combining several components such as: an instrumental model, stellar micro-variability, and in some cases, an additional event, such as a planetary transit, eclipsing binary, or a variable star. The test was blind, as the five different detection teams had no prior knowledge of their content. The main conclusions of this study were:

- The light curves filtering and the removal of systematic effects are almost as important for detecting faint transits as the detection algorithm itself. It never happens that a false event is detected by two independent teams on the same light curve. Therefore, statistical false positives generated by purely random noise are method dependent.
- The BLS algorithm (Kovács et al., 2002) is more efficient to detect faint transits.

The main lesson from these works is that different transit detection algorithms should be used in a transit search.

We have implemented a pipeline with two different transit detection algorithms: the BOX-fitting algorithm (Kovács et al., 2002) and the TRUFAS algorithm (Régulo et al., 2007). They will be used in the LAIWO survey (see chapter 7). In what follows, we describe the implemented algorithms and their improvements. Their efficiency will be studied in the chapters 5 and 6.

4.2 BOX-FITTING algorithm

4.2.1 Overview

The Box-fitting Least Squares (BLS) algorithm (Kovács et al., 2002) is one of the most used and stable techniques for detecting transiting extrasolar planets. It assumes a periodic signal with an alternation between two discrete levels, which represent the in-transit and out-transit parts respectively. A box-shaped profile is fitted to the light curve.

4.2.2 The Method

A strictly periodic signal with period, P_0, is assumed that only takes two different values, H (the out-transit level) and L (the in-transit level). The time spent in-transit is qP_0. The value q is the fractional transit length that is assumed to be small ($\simeq 0.01 - 0.05$), an assumption that is justified because the transiting object is, in general, small in comparison with the star. Thus, the time spent in transit is much less than the time spent out of transit.

The unknown parameters are P_0, q, L, H, and t_0, the epoch of the transit.

Let us call the data set $\{m_i\}$ where $i = 1, 2, \ldots, n$. It is assumed that each $\{m_i\}$ includes a zero-mean Gaussian noise with σ_i standard deviation. The noise is presented by assigning to each data point a weight w_i, defined as:

$$w_i = {\sigma_i}^{-2} \times (\sum_{j=1}^{n} {\sigma_j}^{-2})^{-1}$$

From this definition it is clear that:

$$\sum_{i=1}^{i=n} w_i = 1$$

It is further assumed that the arithmetic averages of the signals $\{m_i\}$ and $\{m_i w_i\}$ are zero.

For any given trial period, every point of the light curve (t_i, m_i) is phase folded and then ordered according to the phase. The phase ϕ_i of any point is given by:

$$\phi_i = (t_i - T_0)/P_t - INT[(t_i - T_0)/P_t]$$

where t_i is the Julian date of the data point, T_0 is an arbitrary epoch, P_t is the trial period and $INT[N]$ denotes the integer part of the number N. By definition $\phi_i \in [0, 1)$.

The folded time series is denoted by $\tilde{m}_i$ with associated weights $\tilde{w}_i$. A step function is fitted to the folded time series with the following parameters: $\hat{L}$, the level in the interval $[i_1, i_2]$ and $\hat{H}$, the level in the intervals $[1, i_1)$ and $(i_2, n]$. The interval $[i_1, i_2]$ corresponds to the transit in the phase folded time series.

The relative time spent at level $\hat{L}$ is given by:

$$r = \sum_{i=i_1}^{i=i_2} \tilde{w}_i \tag{4.1}$$

In other words, the time spent at level $\hat{L}$ with respect to the total time is given by the sum of weights of the data points at the transit level $\hat{L}$.

Given the above definitions, we are searching to minimize the expression (using the same notation as in Kovács et al. 2002) :

$$\mathcal{D} = \sum_{i=1}^{i=i_1-1} \tilde{w}_i(\tilde{m}_i - \hat{H})^2 + \sum_{i=i_1}^{i=i_2} \tilde{w}_i(\tilde{m}_i - \hat{L})^2 + \sum_{i=i_2+1}^{i=n} \tilde{w}_i(\tilde{m}_i - \hat{H})^2 \tag{4.2}$$

Therefore we impose:

$$\frac{\partial \mathcal{D}}{\partial \hat{L}} = 0, \frac{\partial \mathcal{D}}{\partial \hat{H}} = 0 \tag{4.3}$$

The first condition of equation 4.3 gives:

$$\hat{L} = \frac{\sum_{i=i_1}^{i=i_2} \tilde{w}_i \tilde{m}_i}{\sum_{i=i_1}^{i=i_2} \tilde{w}_i}$$

and the second one:

$$\hat{H} = \frac{\sum_{i=1}^{i=i_1-1} \tilde{w}_i \tilde{m}_i + \sum_{i=i_2+1}^{i=n} \tilde{w}_i \tilde{m}_i}{\sum_{i=1}^{i=i_1-1} \tilde{w}_i + \sum_{i=i_2+1}^{i=n} \tilde{w}_i}$$

By using equation 4.1, it is easy to demonstrate that:

$$\sum_{i=1}^{i=i_1-1} \tilde{w}_i + \sum_{i=i_2+1}^{i=n} \tilde{w}_i = 1 - r$$

On the other hand, by assuming $\sum_{i=1}^{i=n} \tilde{w}_i \tilde{m}_i = 0$ (zero arithmetic average) and using the following definition:

$$s \equiv \sum_{i=i_1}^{i=i_2} \tilde{w}_i \tilde{m}_i \tag{4.4}$$

we obtain:

$$\sum_{i=1}^{i=i_1-1} \tilde{w}_i \tilde{m}_i + \sum_{i=i_2+1}^{i=n} \tilde{w}_i \tilde{m}_i = -s$$

Thus, minimization of $\mathcal{D}$ gives:

$$\hat{L} = \frac{s}{r}, \hat{H} = \frac{-s}{1-r} \tag{4.5}$$

In equation 4.5 we see only the weighted arithmetic averages over the proper intervals. Furthermore, if we put 4.5 into equation 4.2 we obtain:

$$\mathcal{D} = \sum_{i=1}^{i=n} \tilde{w}_i {\tilde{m}_i}^2 - \frac{s^2}{r(1-r)} \tag{4.6}$$

If we pay attention to equation 4.6 it is easy to see that the first term on the right hand side does not depend on the trial period or transit interval and, consequently, it is a constant of the light curve. In fact, the minimization of $\mathcal{D}$ is equivalent to the maximization of the second term on the right hand side of equation 4.6, so one can use it alone to characterize the quality of the fit. Therefore, the *Box-fitting Least Squares* (BLS) frequency spectrum is defined as:

$$SR = MAX\left\{\left[\frac{s^2(i_1,i_2)}{r(i_1,i_2)(1-r(i_1,i_2))}\right]^{\frac{1}{2}}\right\} \tag{4.7}$$

If we use the definition of the transit depth, $\delta = L - H$, and the expressions in equation 4.5, we obtain that according to the fit an estimation of the transit depth is:

$$\hat{\delta} = \frac{s}{r(1-r)}$$

Thus, it follows that $SR = \hat{\delta}\sqrt{r(1-r)}$. Thus, we can see that at the correct period, SR yields also an estimate of the transit depth.

Figure 4.1 shows the phase-folded light curve of the confirmed planet OGLE-TR-113, one of the released candidates of the OGLE project (Bouchy et al. 2004, see chapter 5). The BLS fit is superimposed. Figure 4.2 shows the BLS spectrum of OGLE-TR-113. The highest peak is the final period found by BLS. The other peaks are, in general, harmonics or subharmonics of the final period.

In summary, given a trial period, one has to iterate the interval (i_1, i_2) over all points of the folded light curve and find the maximum of SR. Once we have the maximum for each trial period, the absolute maximum is the right true period. Moreover, all the other important quantities are also determined; the transit duration, the mid-transit epoch and transit depth, parameters that are calculated assuming a box-shaped transit.

4.2.3 Signal Detection Efficiency

To characterize the BLS effectiveness, the "Signal Detection Efficiency" is defined as (see also Alcock et al. 2000):

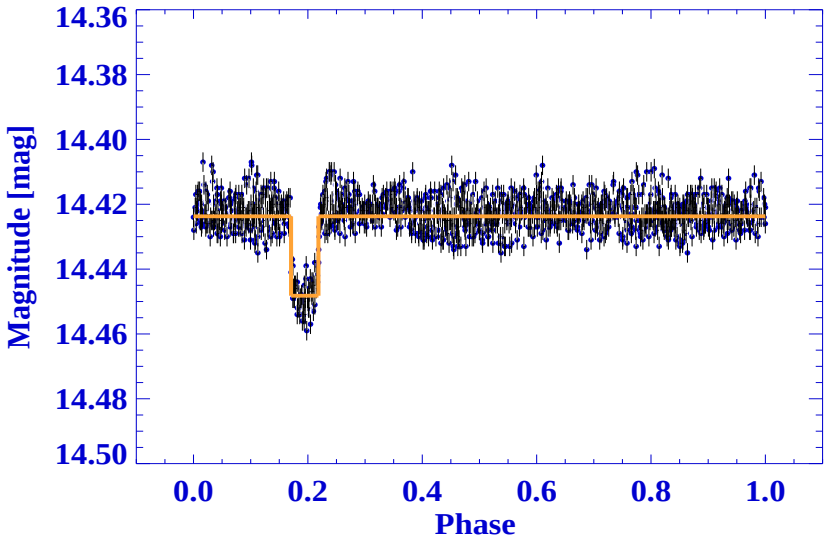

Figure 4.1: Phase-folded light curve of the confirmed planet OGLE-TR-113 (Bouchy et al., 2004). Our BLS fit is superimposed.

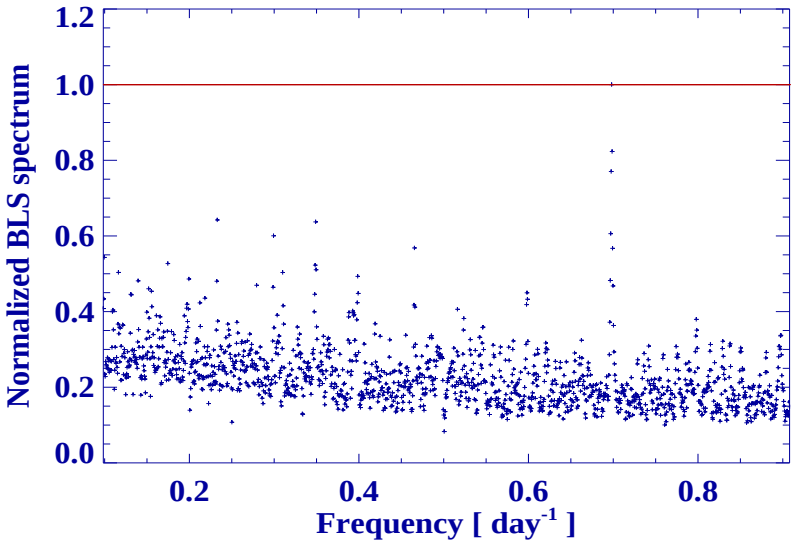

Figure 4.2: Normalized BLS spectrum of the candidate OGLE-TR-113 (Bouchy et al., 2004). The highest peak represents the final period found by BLS.

$$SDE = \frac{SR_{peak} - < SR >}{sd(SR)} \tag{4.8}$$

where SR_{peak} is the BLS spectrum at the highest peak, and $< SR >$ and $sd(SR)$ are the average and standard deviation of SR over the frequency band tested respectively. Simulations by Kovács et al. (2002) demonstrated that a high number of bins yields a high, more stable SDE (see Fig. 3 and Fig. 4 of Kovács et al. 2002). The most important parameter to define a detection as statistically significant is the effective signal-to-noise ratio (α) of the transit:

$$(S/N) \equiv \alpha \equiv \frac{\delta}{\sigma} \times \sqrt{nq} \tag{4.9}$$

where $\sigma/\sqrt{nq}$ is the standard deviation of all measurements within the transit. Here δ is the transit depth, σ is the standard deviation of one measurement, n is the total number of points and q is the fractional transit length.

The region of α around 6 is critical in all cases, because of the separation between the stochastic and deterministic detections (see Fig. 6 of Kovács et al. 2002). Figure 5.7 shows the SDE as a function of S/N for all OGLE release candidates (Udalski et al., 2002a, 2003, 2002b, 2004, 2002c) using our BLS implementation. In practice, a larger S/N value (> 8.0) is used to secure a BLS detection.

4.2.4 Improvements

The originally proposed BLS assumes an homogeneous distribution of points in the phase folded light curve. This assumption is not true if the phase space was not completely covered. To solve this possible problem, the binning was done based on phase intervals rather than number of points.

The second improvement was the implementation of the "directional correction" of Tingley (2003). This modification takes into account the sign information of the summation that is carried out in the calculation of the BLS statistic (in the transit interval $[i_1, i_2]$), which is otherwise simply squared and therefore lost. A periodic increase in magnitude has the same test statistic as a periodic decrease. This can be corrected by simply not calculating the test statistic for any test transit where the weighted sum of the in-transit differential magnitudes is negative (and therefore has increased in brightness). The weighted summation (using 4.1 and 4.4) in the transit interval is defined as:

$$D \equiv \frac{s}{r}$$

Thus, if $D < 0$ the BLS statistic is not calculated, because it is a brightness increase rather than a transit (the magnitude scale is inverse).

4.3 TRUFAS algorithm

4.3.1 Overview

The TRUFAS algorithm (Régulo et al., 2007) is a transit detection algorithm based on wavelet techniques. It is designed for data with continuous time coverage, like the space mission CoRoT (Baglin, 2003).

The procedure works basically in two steps: a continuous wavelet transformation of the light curve with a posterior selection of the optimum scale for transit detection, and a period search in the power spectrum of the selected wavelet transformation.

4.3.2 The Method

A continuous wavelet transform (CWT) of a function $f(\eta)$ is defined by:

$$CWT \equiv \int f(\eta)\Psi^*(\eta)d\eta \tag{4.10}$$

The function $\Psi(\eta)$ is called the "mother" wavelet, because it is the base of the wavelet transform, i.e., the transformation is done using shifts and scales of this function. The wavelet transform is basically a convolution of the light curve with the wavelet function. Moreover, the idea is to amplify the transit signal with this convolution. The Paul Function (Torrence & Compo 1998) of order 1 is chosen as the mother wavelet because its shape is similar to the feature we are looking for, i.e., a transit. The Paul Function of order "m" is defined as:

$$\Psi_0(\eta) = \frac{2^m i^m}{\sqrt{\pi(2m)!}} \times (1 - i\eta)^{-(m+1)}$$

and the scaled displaced wavelet is:

$$\Psi\left(\frac{t - t_0}{s}\right) = \left(\frac{1}{s}\right)^{1/2} \Psi_0\left(\frac{t - t_0}{s}\right)$$

Here t is the time, s is the dilation parameter used to change the scale, and t_0 is the translation parameter used to slide in time. The factor $s^{-1/2}$ is a normalization factor to keep the total energy of the scaled wavelet constant. This normalization is done to allow the comparison of different wavelet transformations. Figure 4.3 shows an example of the scaled Paul wavelet

function of order 1. The Gaussian-like part of the function corresponds to a certain transit duration ($\sim$ 1.2 hours in this case).

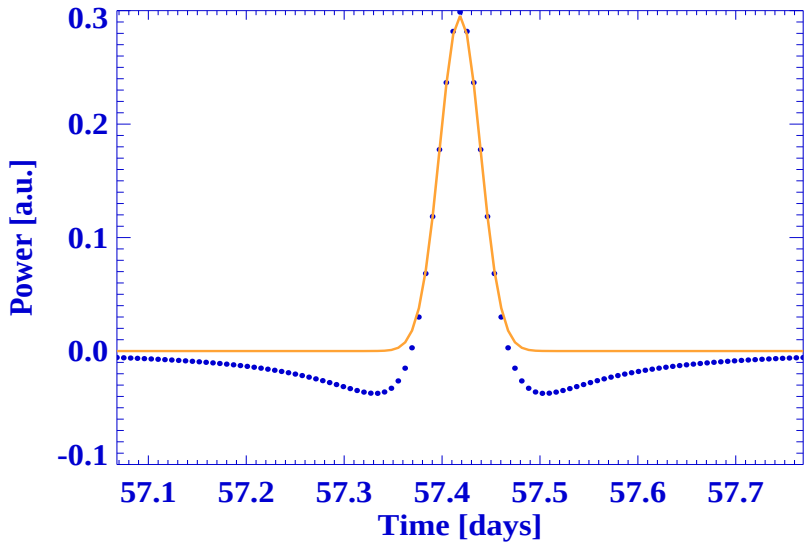

Figure 4.3: Imaginary part of Paul wavelet of order 1. The dotted curve shows the function and the solid curve shows a Gaussian fit to it. In this case, the scale corresponds to a transit duration of 1.2 hours (the FWHM of the Gaussian fit).

The light curve is convolved with 55 different scales that correspond to transit durations ranging from $\sim$ 1.5 hours to $\sim$ 50 hours. The higher the correlation, the higher the coefficients of the CWT. The selection is done using a double criteria: the scale with higher coefficients is selected when these coefficients are present in more scales. This is done because higher coefficients may be produced by random noise. Therefore, the idea is to look at many scales to probe that they are produced by a real signal, since real signals must appear in other scales too. If the automatic scale selection fails (due to low S/N transits), a scale that corresponds to a transit duration of 5.7 hours is used. All the former parameters (range of scales, number of scales or default scale) can be changed to be more suitable for ground-based transit planet searches.

If there is a clear transit signal in the light curve, it will appear as equally spaced peaks in the selected scale, as is shown in Figure 4.4 for the light curve #533 of the "CoRoT Blind test 1" (Moutou et al., 2005).

The case of some ground-based planet searches (like the OGLE survey, see chapter 5) is

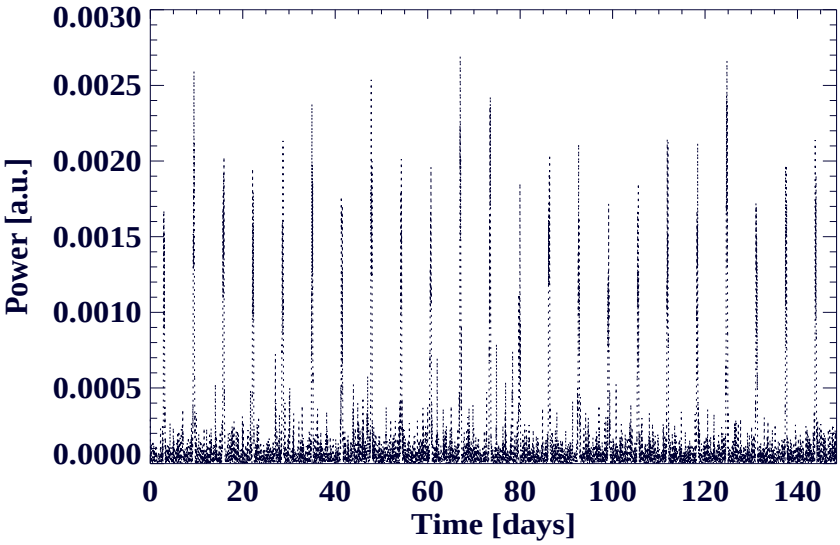

Figure 4.4: Square Wavelet of the selected scale for light curve #533 of the set of simulated light curves of (using our implementation, Moutou et al. 2005). The units of the vertical axis ("power") are the square of the signal of the CWT.

slightly more complicated. If the sampling is not constant, gaps will appear in the square wavelet transformation. This can be seen in figure 4.5 for the confirmed planet OGLE-TR-113 (Bouchy et al., 2004).

The next step of TRUFAS is based on the method developed in Régulo & Roca Cortés (2002). The power spectrum of the wavelet square of the selected scale is calculated. This spectrum is again a series of equally spaced peaks, but now the first peak is at zero frequency, independent of the epoch of the transits (see Figures 4.6 and 4.7). Knowing the position of the first one, finding the spacing among the peaks is now much easier. The search for periodicity is done iteratively trying a range of periods between the expected values. The idea is to find a signal 1.5 times above the RMS of the power spectrum in the searched periods. To avoid binning effects and to evaluate the significance of the detection, this procedure is repeated 50 times on the selected scale, but continuously shortening its length, until it is shortened to about 10%.

4.3.3 Significance of the detection

To measure the significance of the detection, and therefore recognize false detections due to random noise, the signal is reconstructed. Since we know the period, we can select only the

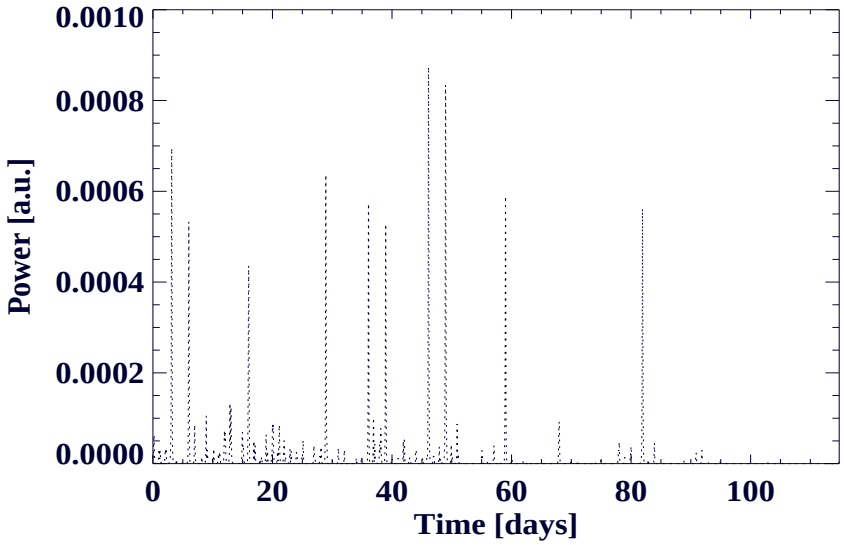

Figure 4.5: Square wavelet of the selected scale for the confirmed planet OGLE-TR-113 (Bouchy et al., 2004). The gaps of the observations are clearly a disadvantage to recognize the true period.

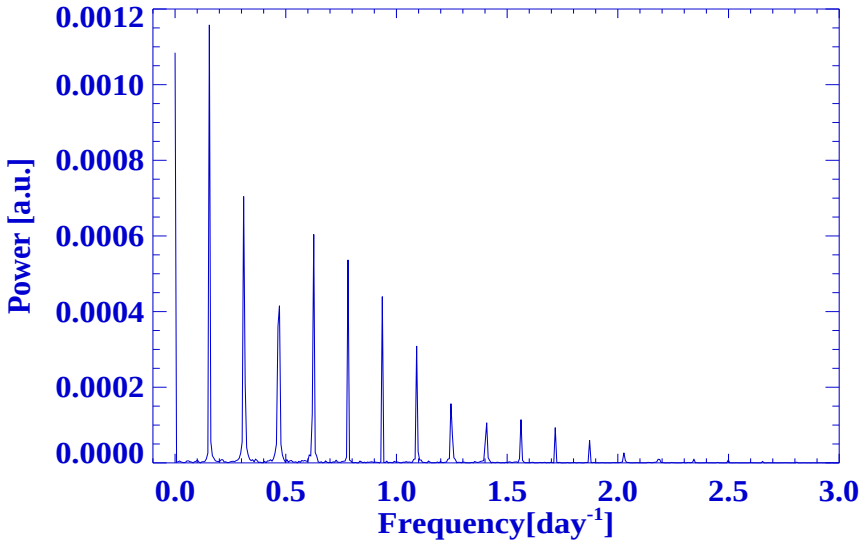

Figure 4.6: Power spectrum of the wavelet square for the selected scale of light curve #533. If there is a periodic signal, regularly spaced peaks will appear. First peak is at zero frequency.

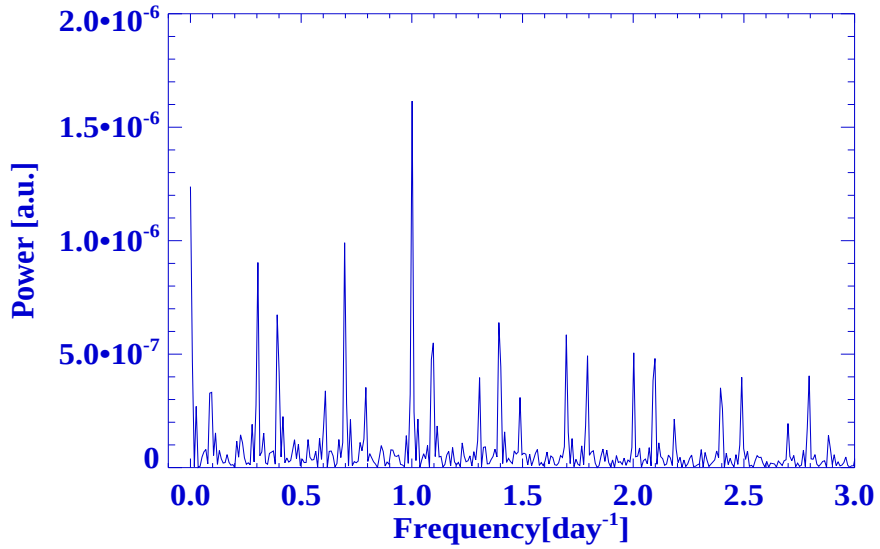

Figure 4.7: Power spectrum of the wavelet square for the selected scale of the confirmed planet OGLE-TR-113.

complex Fourier components corresponding to signal and perform an inverse Fourier transform. If the detection is a real transit, the result will be a signal with a much better S/N, as can be seen in Figure 4.8. If the amplitude (A) of the recovered signal is compared with the sigma (σ) of the selected scale, the ratio A/σ defines a threshold of higher than 1 for real transits and less than 1 for false detections.

4.3.4 Improvements

TRUFAS is tailored for continuous observations, like ground-based network surveys or space-based observations, where the time sampling usually is constant. In ground-based observations with gaps (like the OGLE survey), where the time coverage is not complete, this is not the case. For these type of observations, it is necessary to redistribute the information in order make it homogeneous in time space. To accomplish this a binning process was performed, with the precaution of including at least 2 observations within a minimum transit duration interval, that was chosen to be $\sim$ 2 hours. The latter constraint always imposes a minimum number of bins. In addition, since in some cases the required number of bins is big, and therefore the number of data points[1] is big, another constraint was fulfilled: the number of bins must be a multiple

[1]If there is no information in a bin, the value is set to zero.

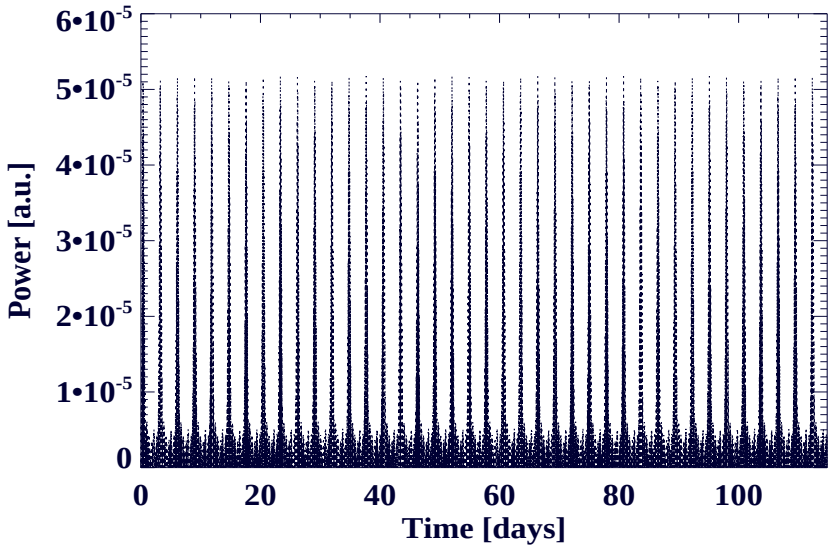

Figure 4.8: Recovered wavelet square of the confirmed planet OGLE-TR-113. The reconstructed wavelet square gives a signal with a much higher signal to noise.

of two, because this allows a fast Fourier transform (FFT) to be used to calculate the wavelet transform (Torrence & Compo 1998). The use of FFT speeds up the calculation.

Ground-based observations have gaps, so in some cases just part of the transit is observed. The selection of the scale requires the presence of individual transits. Therefore a missing part of a transit will cause this process to be not as robust as desired. Thus, the possibility to fix the scale to a reasonable value was also included, being a good compromise for the expected length of transits.

5

The OGLE data

5.1 The OGLE Project

The Optical Gravitational Lensing Experiment (OGLE) (Udalski et al., 2002a, 2003, 2002b, 2004, 2002c) was an extensive photometric search for planetary and low-luminosity object transits in Galactic disk stars (third phase called OGLE-III). Observations in the I-band filter were collected using the 1.3-m Warsaw telescope at Las Campanas Observatory, Chile, equipped with a CCD camera with a field of view of 35' x 35'.

The OGLE team released a list of 177 low-luminosity transit candidates, out of which a handful turned out to be exoplanets: OGLE-TR-56b (Konacki et al. 2003, the first extrasolar planet discovered primarily by the transit method); OGLE-TR-113b and OGLE-TR-132b (Bouchy et al., 2004); OGLE-TR-111b (Pont et al., 2004) and OGLE-TR-10b (Konacki et al., 2005). Moreover, new analysis of OGLE data has delivered additional transiting planets: OGLE-TR-182 (Pont et al., 2008), OGLE-TR-211 (Udalski et al., 2008) and OGLE2-TR-L9b (Snellen et al., 2009). Table 5.1 summarizes the main properties of the released OGLE planets.

5.2 The OGLE releases

In the first release (Udalski et al., 2002a), three fields in the direction of the Galactic center were observed. No robust transit detection algorithm was known and the sample was selected by the presence of apparent transits only, with no knowledge on any properties of the systems, like mass. Moreover, the transiting objects could have been Jupiter-like planets, brown dwarfs, or M dwarfs (they have similar sizes: 0.1 - 0.2 $R_{\odot}$). A measurement of the radial velocity amplitude of the stars is always needed to determine the masses of the transiting companions. In this release

Table 5.1: Parameters of the OGLE planets

OGLE Candidate	Period [day]	Transit Duration [hour]	Transit Depth [mag]
OGLE-TR-10b	3.10129	2.2	0.019
OGLE-TR-56b	1.21191	1.5	0.013
OGLE-TR-111b	4.01445	2.3	0.019
OGLE-TR-113b	1.43248	1.5	0.030
OGLE-TR-132b	1.68987	2.0	0.011

- Periods taken from exoplanet.eu.
- Transit durations taken from Tingley & Sackett (2005).
- Transit depths taken from OGLE releases.

there was a first selection of the candidates based on a color-magnitude diagram (to separate main sequence disk stars from subgiants and giants; this was done because the observed fields were in the direction of the Galactic center). In addition, only stars with photometry precision $\leq$ 0.015 mag (1.5%) were kept. The selection of the transit candidates was done using an error-less transit light curve with an amplitude of 0.015 mag and a total duration of 0.03 in phase space. All the observations of a particular light curve were phase-folded with trial periods between 1 and 10 days (with $10^{-4}\times$ P as step, where P is the period), and then they were cross-correlated with the artificial light curve. If the cross-correlation coefficient was larger than a preselected threshold the star was marked as a candidate. The final check of candidates was done by a careful visual inspection of the photometric data, both in time and phase space. The final periods of the candidates were found by an exhaustive examination of the eclipse light curve, by minimizing dispersion in the eclipse phases. Forty six transit candidates (IDs from 1 to 46) were found in this release, after selecting transit depths smaller than 0.08 mag (corresponding to 1.4 R_{Jup} if the stellar radius is $0.5R_{\odot}$).

The second release (Udalski et al., 2002c) was a supplement of the first release, because the OGLE team was encouraged by the new transit search technique: the BOX-fitting Least Square method (BLS, see chapter 4). They ran BLS on $\sim$ 52,000 stars selected for transit search in Udalski et al. (2002a). All of the proposed candidates of the first release were easily found by BLS and, in addition, 13 new candidates were detected (IDs 47 to 59). BLS was run for all the data using the following parameters: 200 phase bins, frequencies from 0.1 $[day^{-1}]$ (P = 10 [days]) to 0.95 $[day^{-1}]$ (P = 1.053 [days]) with a frequency step of 10^{-4} $[day^{-1}]$ and fractional transit

length (transit duration divided by period) between 0.015 and 0.15 in phase space. Only objects with S/N greater than 9 and SDE greater than 3 (see definitions in 4.9 and 4.8) were considered further. The lowest S/N detections were also required to have a proportionally larger SDE value to avoid too many false detections. A final selection cut was done by looking at the light curves for clear triangle-shaped eclipses (most likely grazing eclipses of stars) or objects with secondary eclipses (very likely eclipsing stars blended with brighter stars). The first transiting extrasolar planet discovered first by the transit method (OGLE-TR-56, Konacki et al. 2003) was proposed in this supplement.

In the third release (Udalski et al., 2002b), three fields in the direction of Carina ($l \simeq 290$ deg) were observed. The reason to choose other fields away from the direction of the Galactic Center was to reduce the blending. It was also decided to increase the exposure time (180 seconds) compared with previous campaigns (120 seconds) at the cost of reducing the time sampling (from 12 min to 15 min) in order to reach fainter stars, i.e., in general later spectral types. Moreover, fields located away from the direction of the Galactic Center reduce the contamination of giants, so no color-cut was necessary. The BLS algorithm was again run and the final list of candidates was prepared after a careful visual inspection. This was done to eliminate V-shaped transits most likely caused by grazing eclipses of stars and to eliminate deeper transits that are, in general, not caused by planets. Even though the searched periods were from 1.05 - 10 days, it was possible to find planets with smaller periods by finding their harmonics. 62 new candidates (IDs 60 to 121) were proposed in this release.

The fourth release (Udalski et al., 2003) was a supplement of the previous releases. Small-scale systematic effects started to be recognized (even visual inspection of the data indicates that they are present) and there was the possibility to remove them using the proposed method of Kruszewski & Semeniuk (2003) (see chapter 3). After removing these effects, 16 additional candidates were found, lost in the noise in previous searches.

In the fifth release (Udalski et al., 2004), six Galactic disk fields were monitored. The usual photometry precision cut (15mmag) was performed. The photometric data of all objects were corrected for small scale systematics (Kruszewski & Semeniuk, 2003). Moreover, the experience of previous campaigns indicated that the size of detectable exoplanets is around Jupiter-like. Radial velocity measurements indicated that none of the transiting planets with deep transits turned out to be a planet, therefore a tighter limit on the depth of transits was used, so all objects with transits deeper than 0.05 mag were removed. As usual, V-shaped transits were also checked and removed. Cases in which a small amplitude sinusoidal variation caused by distortion

of the primary, a clear sign of a relatively massive companion (Drake, 2003; Sirko & Paczyński, 2003), were also removed. This campaign was focused mainly on the smallest companions which have the larger probability to be extrasolar planets. Forty new candidates were found in this campaign (IDs 138 to 177).

Table 5.2 summarizes the OGLE releases. Of the 177 released candidates, there are four candidates (IDs 43 to 46) in which just one transit was observed, therefore no period or other parameters could be determined. They will not be considered here forth in the analysis.

Table 5.2: Summary of OGLE releases.

RELEASE	IDS	No POINTS	Precision [%]	No of fields	Direction	Notes
1.	1 - 46	800	1.5	3	Bulge[1]	
2.	47 - 59	900	1.5	-	-	BLS[2]
3.	60 - 121	1150	1.5	3	Carina (l $\sim$ 290deg)	
4.	122 - 137	1090	1.5	-	-	Systematics[3]
5.	138 - 177	1100	1.5	6	Galactic Disk	

[1] It denotes the direction of the Galactic Center.

[2] BLS detected additional 13 candidates from the first release.

[3] Systematics in previous data were corrected and new candidates were found.

5.3 Testing the transit detection algorithms

The OGLE data is an appropriate set to test our implemented transit detection algorithms (TDA). It contains different types of transiting object light curves, which makes it ideal to test the efficiency of our implemented TDA under different circumstances. The OGLE team determined all the parameters of their released candidates, thus, we can compare them with the ones of our implementations. In what follows, we will describe the results and tests with the aim of establishing the efficiency of our TDA based on different light curve parameters.

5.3.1 Box-Fitting detection efficiency

We ran our BLS with the same input parameters that the OGLE team used: 200 phase bins, frequencies from 0.1 $[day^{-1}]$ (Period = 10 [day]) to 0.95 $[day^{-1}]$ (Period = 1.053 [day]), frequency step of 10^{-4} $[day^{-1}]$ and fractional transit length (transit duration divided by period) from 0.015 to 0.15 in phase space (see Defaÿ et al. 2001).

Figure 5.1 summarizes our results. The top panel shows the difference between the period obtained by the OGLE team (P_{OGLE}) and our results (P_{BLS}). The circles are the clear detections where the difference is less than 1%. A cross plus diamond indicates an harmonic or subharmonic.

A detection is defined as an harmonic (subharmonic) when the ratio (inverse ratio) between the two periods is close to an integer number. Mathematically, a candidate is an harmonic or subharmonic when $R < 0.01$, with R defined as:

$$R = (P_{OGLE}/P_{BLS}) - INT(P_{OGLE}/P_{BLS})$$

for an harmonic, or:

$$R = (P_{BLS}/P_{OGLE}) - INT(P_{BLS}/P_{OGLE})$$

for a subharmonic. Here INT(N) denotes the closest integer to the number N.

Of the 173 candidates, 153 periods were within 1% of the OGLE results (88.4%). All of the OGLE planets were detected. Table 5.3 shows the results obtained by our BLS for the OGLE planets, results that are in good agreement (see Table 5.1) within 0.1%. The last two columns indicate the transit depth and duration that are found if a box-like curve is fitted to the phase folded light curve.

Table 5.3: BLS results for the OGLE planets

OGLE Candidate	Period BLS [day]	Transit Duration BOX [hour]	Transit Depth BOX [mag]
OGLE-TR-10b	3.10130	2.2	0.018
OGLE-TR-56b	1.21192	1.5	0.013
OGLE-TR-111b	4.01598	1.9	0.016
OGLE-TR-113b	1.43254	1.5	0.024
OGLE-TR-132b	1.68973	1.8	0.008

Of the non-detections within 1%, 14 are harmonics or subharmonics, that can be also taken as detections. This raises the number of detections to 167 (96.5%).

Harmonics or subharmonics are inevitable in some cases, because of noisy light curves or gaps in the observations. Figure 5.2 shows the candidate OGLE-TR-5, which is detected as harmonic (its period is 0.8082 [day], and the period found by BLS is 1.6179 [day]). The period of this candidate is outside the searched range, but it is detected as an harmonic, like it happened for the OGLE team in some cases too. It is important to mention that the BLS statistic (see Figure

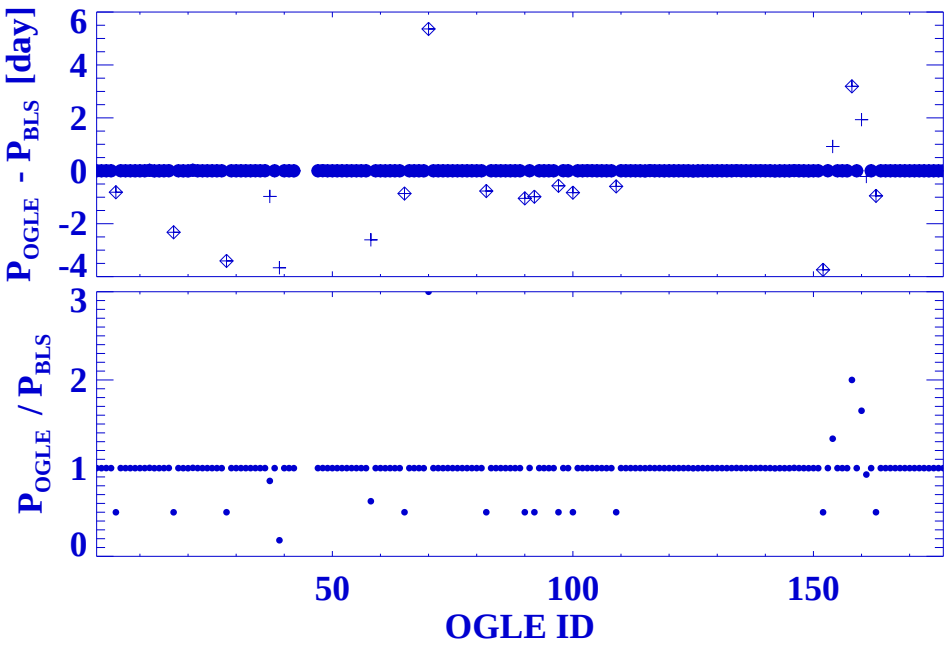

Figure 5.1: Comparison of the period determined by the OGLE team and our results. In the top panel the difference is plotted. The circles are clear detections (difference is less than 1%). Cross plus diamond denotes an harmonic or subharmonic (for definition see text). A cross (in the top panel) indicates a non-detection. The bottom panel shows the ratio of period determined by OGLE and our results.

5.3) is calculated only in the transit interval (where the BOX is in Figure 5.2) and not using the complete phase folded light curve, which is why harmonics or subharmonics are sometimes detected.

Figure 5.4 shows the phase folded light curve of the planet OGLE-TR-56, where a clear detection is present.

The BLS spectrum (see 4.2) has several peaks (see Figure 5.3 for an example), so it is possible to detect the real period of a planet in the second or even in the third most significant peak (this often happens in noisy light curves). If we check the second and third peaks of the BLS spectrum of the non-detected candidates, we find four more candidates (i.e. those peaks correspond to the true period or harmonics). The latter increases the positive detections to 171 (98.8%).

Figure 5.5 shows the phase folded light curve (wrong period) of the candidate OGLE-TR-161,

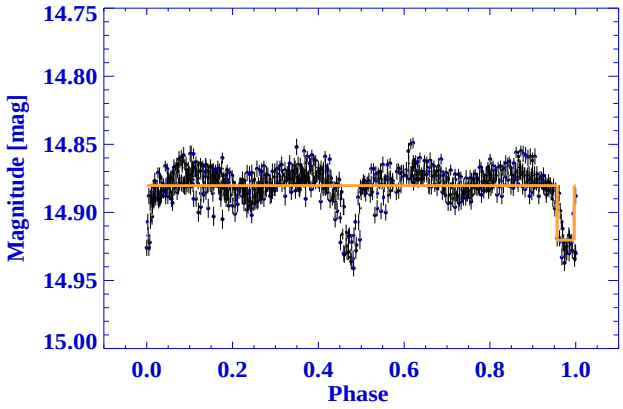

Figure 5.2: Phase folded light curve of OGLE-TR-5. Our BLS fit is superimposed. Light curve where clearly an harmonic was detected.

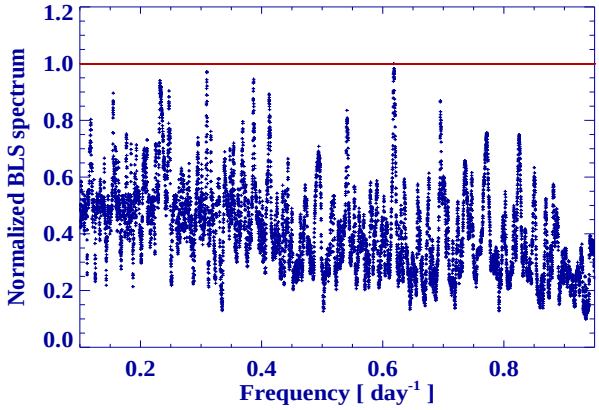

Figure 5.3: Normalized BLS spectrum for the candidate OGLE-TR-5. The maximum peak (where the red line is) is the final period found by BLS. It is common to have several peaks that are harmonics or subharmonics of the real period.

as an example where BLS fails.

Figure 5.6 shows the real light curves of the candidate OGLE-TR-161 in time and phase space (correct period). Just visual inspection shows that the transit is faint.

Figure 5.7 (top) shows the Signal detection efficiency (SDE) for the 173 OGLE candidates calculated using our BLS implementation (see definition in 4.2). The median of all SDE is 6.0. There is no candidate with SDE lower than 3.0, a number that was used as a lower limit to

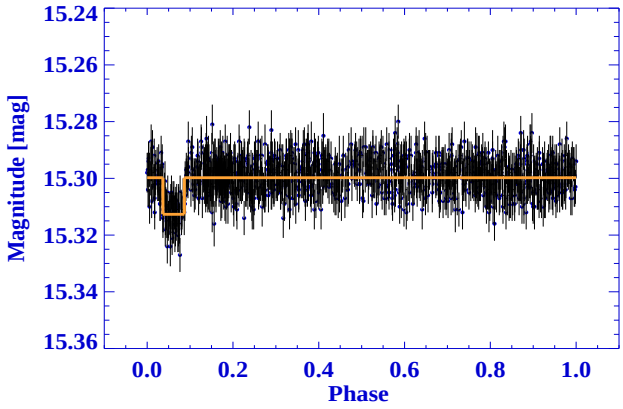

Figure 5.4: Phase folded light curve of OGLE-TR-56. Our BLS fit is superimposed.

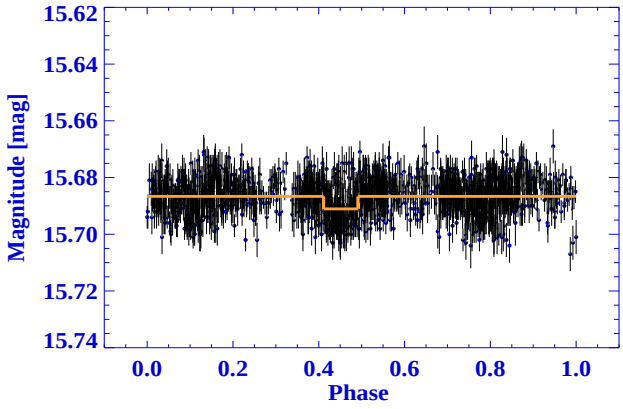

Figure 5.5: Phase folded light curve of OGLE-TR-161. The phase is calculated with a wrong period determined by BLS. Our BLS fit is superimposed.

define a detection for the OGLE team. There is no candidate with S/N lower than 6.0. Figure 5.7 (bottom) shows the same plot, but now as a function of the S/N including red noise (see equation 3.6). When we include red noise, we are taking into account correlated noise between points in the transit. This has the effect of lowering the S/N in the case of candidates with big systematic effects, especially for the bright stars. The values of S/N, including red noise, were obtained assuming a red noise value (σ_r) of 4 mmag (see section 3.2).

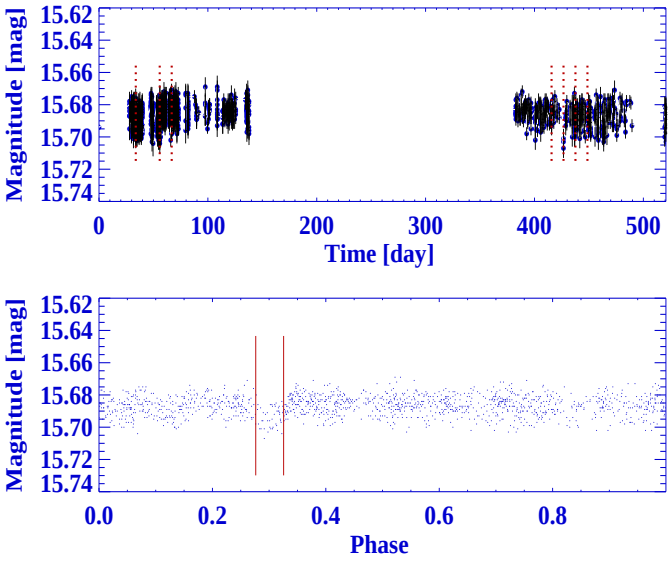

Figure 5.6: Light curves of OGLE-TR-161. Top panel shows the light curve in time space. The dotted lines indicate the position of individual transits. Bottom panel shows the phase-folded light curve with the correct period determined by the OGLE team, the lines indicate the position of the transit.

Consequently, almost all the OGLE candidates are securely detected by our BLS according to the OGLE criteria.

Finally, figure 5.8 compares the estimated transit durations and depths for the positive detections. The transit durations were taken from Tingley & Sackett (2005). They were estimated using a simple matched filter code, so they are strictly just reference values and not directly comparable with the values obtained using a BOX-shaped transit. In addition, a BOX-shaped transit does not take into account the time of ingress and egress. The transit durations estimated by Tingley & Sackett (2005) (D_{OGLE}) are systematically higher than the ones estimated using our BLS (D_{BLS}), with an average deviation of 13.6%. For the transit depths there is good agreement too (OGLE values are in average 13.8% higher). This is in part explained because a BOX-shaped transit does not consider limb-darkening.

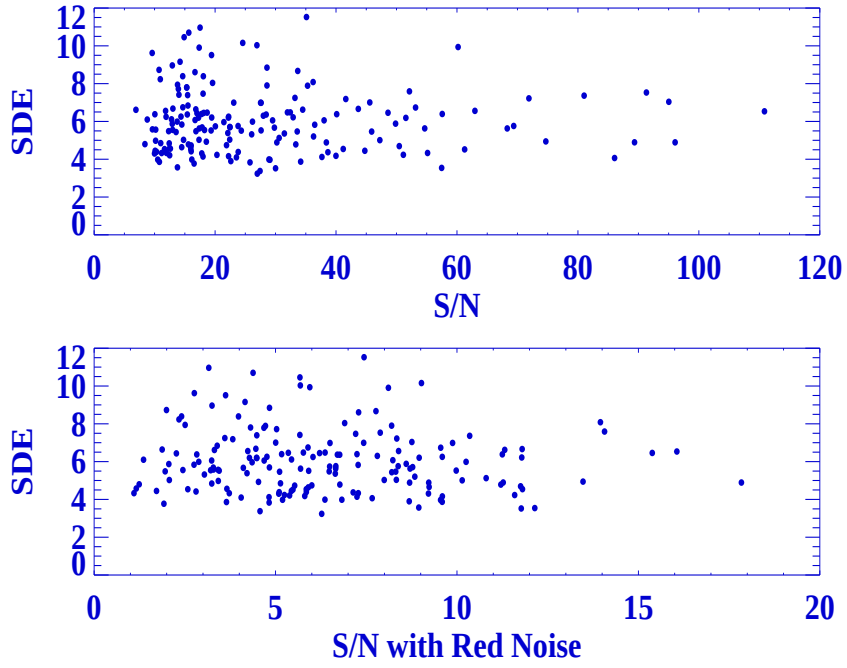

Figure 5.7: SDE as a function of S/N determined using our implementation of BLS. In the upper panel only white noise is considered. In the lower panel the red noise is taken into account.

5.3.1.1 Detection Efficiency

There are 2 non-detections (OGLE-TR-160, 161) in which our BLS did not recognize the true period reported by the OGLE team. It is now possible to study the efficiency of our implemented algorithm given several light curve parameters, and therefore recognize its weaknesses based on data quality.

Figure 5.9 shows the BLS detection efficiency as a function of number of points in transit. No dependence in the number of points in transit is observed. Thus, it is not a limiting factor of BLS for OGLE type observations.

Figure 5.10 shows the BLS detection efficiency as a function of number of transits present in the light curve. Again, there is no clear dependence between BLS efficiency and number of transits for OGLE light curves. Thus, the number of transits does not challenge BLS for OGLE type ground-based observations (i.e. number of points greater than 800 and time sampling between 12 and 15 min).

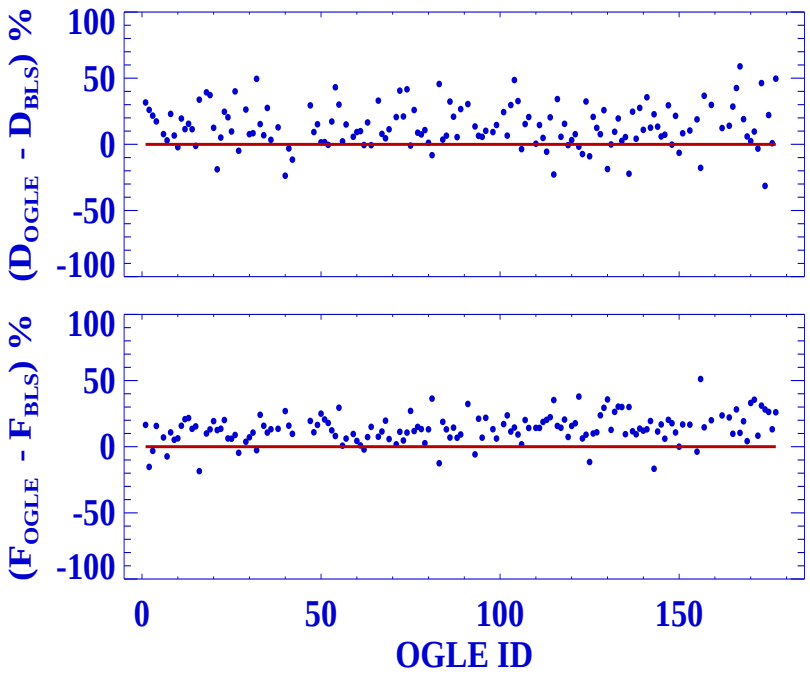

Figure 5.8: Comparison of the estimated transit durations (top) and transit depths (bottom) with values taken from Tingley & Sackett (2005) and from the OGLE team values. The red line is the zero value as a reference.

Figure 5.11 shows the BLS detection efficiency as a function of transit depth. Clearly, the transit depth influences the BLS detections; the efficiency decreases to 75% when the transit depth is below 0.75%. So, shallower transits challenge our BLS implementation.

A key parameter that includes almost all light curve parameters is the signal-to-noise ratio (S/N) of the transit (see equation 4.9). Figure 5.12 shows the BLS detection efficiency as a function of S/N. Even though the S/N includes the transit depth, it does not show the same strong dependence. This is, in part, explained because more points decrease the noise of the transit, thus increasing the S/N.

Finally, figure 5.13 shows the BLS efficiency as a function of S/N that takes into account the red noise (see equation 3.6). We can clearly see that it is one of the most clearest indicators of the BLS detection failures. The BLS efficiency decreases to *66.7%* when the S/N including red noise is below *5.1*.

In conclusion, the transit depth and correlated noise are the main limiting factors that

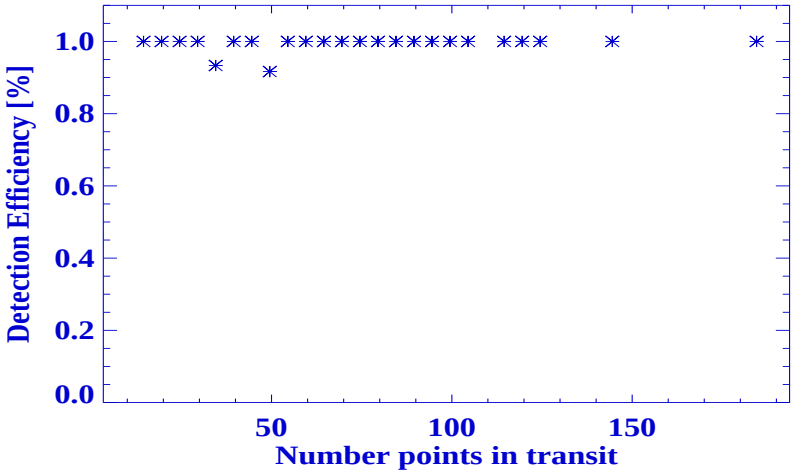

Figure 5.9: BLS detection efficiency as a function of number of points in transit.

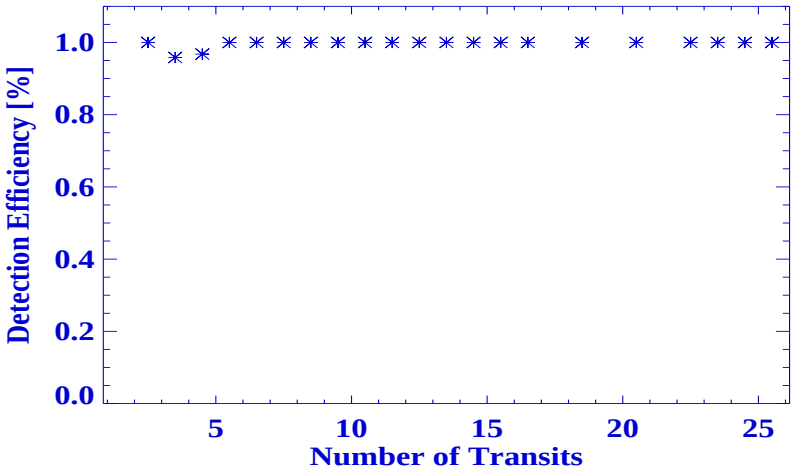

Figure 5.10: BLS detection efficiency as a function of number of transits.

challenge BLS for the OGLE-type ground-based observations.

5.3.1.2 Execution time

Since BLS is often executed for a large set of light curves, it is important to check the speed of the implemented algorithm as a function of two important parameters: the number of bins in the phase folded light curve and the number of trial frequencies to search for periods. We

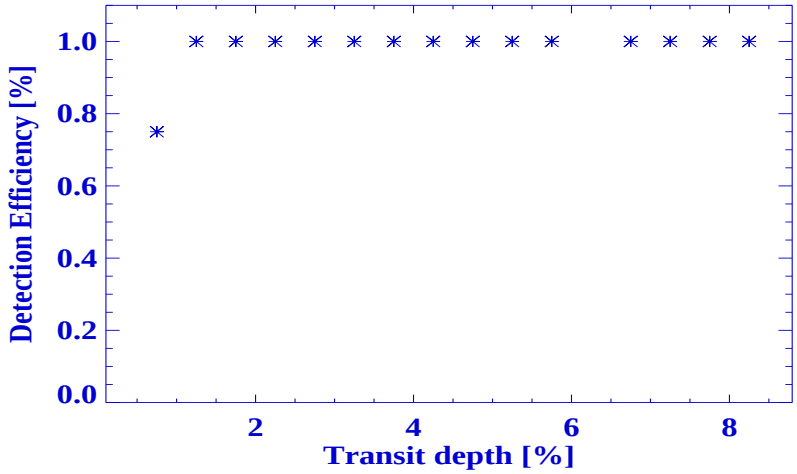

Figure 5.11: BLS detection efficiency as a function of transit depth.

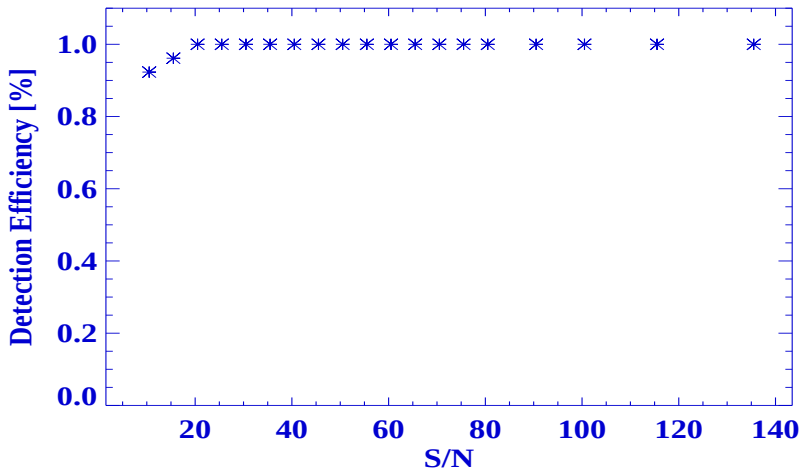

Figure 5.12: BLS detection efficiency as a function of S/N.

applied 7 times the BOX-fitting algorithm to the 173 OGLE candidates for each test, varying the number of bins while keeping the number of frequencies constant and vice versa. The number of frequencies was set constant to a value of 500 when testing the number of bins. The number of bins was set constant to a value of 500 when testing the number of frequencies. Figure 5.14 summarizes the results. As we can see, the execution time varies more rapidly with the number of bins than with the number of frequencies. This is in part explained because the number

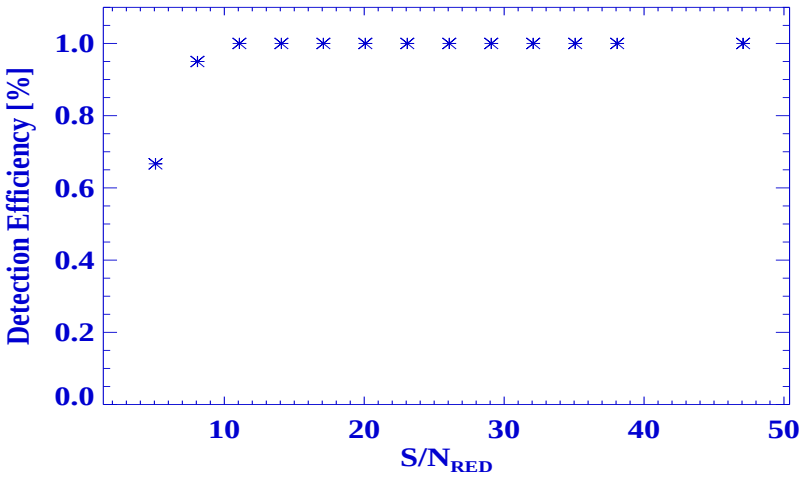

Figure 5.13: BLS detection efficiency as a function of S/N with red noise.

of operations in the BLS algorithm is proportional to $m \times (\Delta_{MAX} - \Delta_{MIN})$, where m is the number of bins and Δ_{MAX} and Δ_{MIN} are the maximum and minimum fractional transit length allowed (see section 3.1 in Kovács et al. 2002).

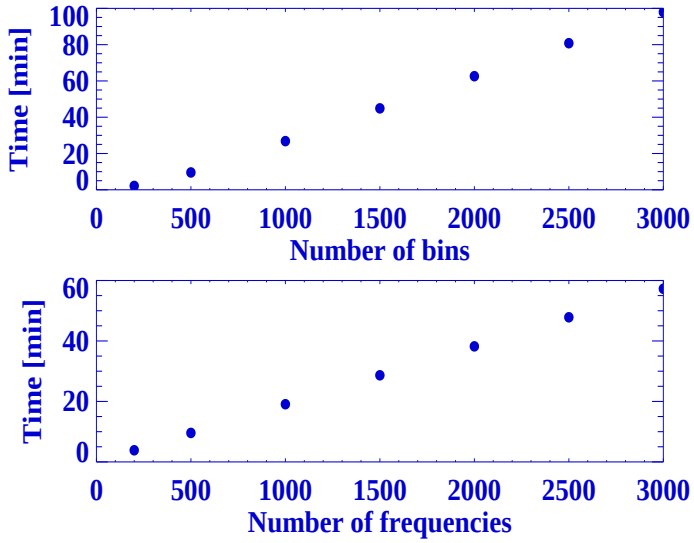

Figure 5.14: Results of two different test of the speed of the implemented BOX-fitting algorithm. The execution time varies more rapidly with number of bins than with the number of frequencies.

5.3.1.3 Light Curve Degradation

Time Sampling

A good time resolution ensures a good sampling of the light curve. There are cases, however, when the time sampling has to be degraded (to reach fainter targets for example). We continuously degraded the time sampling of all OGLE candidates and ran BLS in every iteration to test how different time samplings challenge it. In every iteration, the minimum time sampling was determined, and then points that were at that minimum time sampling were removed. In the next iteration, the minimum time sampling was made bigger. Thus, more points were iteratively removed.

Table 5.4 shows the first two iterations for example candidates (IDs 5, 37, 88, 126, and 175). In most of the cases the first degradation was the most important one, in the sense that the total number of points decreased to half of the initial value. In the following iterations, the total number of points decreased slowly. Therefore, the BLS results did not show a big change and remained stable. In general, it is observed that the BLS results are robust when the time sampling changes, because the efficiency does not decrease. This can be explained because the BLS statistic is calculated in the phase-folded light curve and the S/N of the transit decreases only by a factor $\sim 1/\sqrt{2}$ when the number of points in transit decreases to half.

Table 5.4: Time sampling degradation results for some candidates

OGLE ID	Time Sampling $[min]$	Total Points	Real Period $[day]$	P_{BLS} 1 $[day]$	P_{BLS} 2 $[day]$	P_{BLS} 3 $[day]$
5	8.9	808	0.80820	1.61772	1.61624	4.31136
	12.3	461	0.80820	4.31136	1.61624	6.45453
37	8.6	801	5.71970	5.74106	5.72246	6.69940
	12.3	460	5.71970	5.72246	5.74106	6.69940
88	10.3	1131	1.25010	1.25010	1.25099	1.24922
	15.6	637	1.25010	1.25010	1.25099	1.24922
126	10.3	1153	5.11080	5.10979	5.12462	5.08039
	14.5	658	5.11080	2.55707	5.10979	5.12462
175	11.1	1349	1.48830	1.48820	2.97995	3.02073
	14.5	788	1.48830	1.48820	3.02073	3.05205

- P_{BLS} X, denotes the period found in the X peak of the BLS spectrum.

Removals of individual points

To strongly challenge BLS another test was run. We degraded all the OGLE light curves by removing every second point. In this way, it is possible to remove half of the points in every iteration, and therefore, to change the parameters (total number of points, points in transit, S/N of the transit, etc.) more rapidly. Table 5.5 shows the average parameter values of the degraded light curves after every iteration. As was expected, the parameters change faster from one iteration to another. Figure 5.15 shows the BLS detection efficiency as a function of every iteration. As can be seen, the BLS results are robust, because its detection efficiency does not decrease much when the parameters experience a big change.

Table 5.5: Average light curve parameter values per iteration

ITERATION	TOTAL POINTS	POINTS IN TRANSIT	S/N	TIME SAMPLING [min]
0	1026.7 +/- 227.5	49.2 +/- 27.5	36.13 +/- 25.57	14.6 +/- 1.4
1	513.6 +/- 113.8	24.7 +/- 13.9	25.58 +/- 18.33	29.7 +/- 2.7
2	257.0 +/- 56.9	12.2 +/- 7.1	17.92 +/- 12.80	66.4 +/- 9.1
3	128.7 +/- 28.5	6.2 +/- 3.9	12.64 +/- 9.14	545.5 +/- 510.0

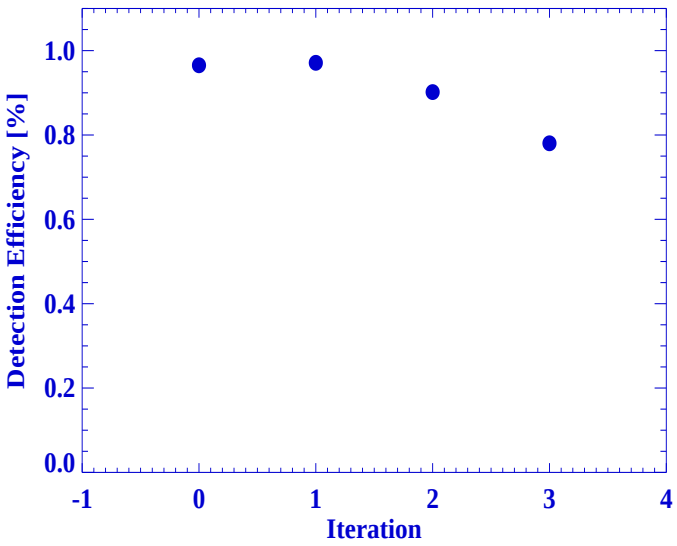

Figure 5.15: BLS detection efficiency as a function of degradation iteration.

5.3.2 TRUFAS detection efficiency

We ran TRUFAS using the following parameters: 55 scales that correspond to transit durations from 1.5 to 10.0 hours, periods from 0.9 to 10.1 days, and thresholds of 1.5 times the RMS in the power spectrum of the wavelet transform. The selected scale was shortened to 10% in 50 iterations. The former parameters are the recommendation of Régulo et al. (2007). In addition, we required to have at least two points in an interval of 1.5 hours (when binning the light curve in time).

Figure 5.16 shows the comparison between the period found by TRUFAS and the period reported by the OGLE team. The top panel shows the difference between the period obtained by the OGLE team (P_{OGLE}) and our results (P_{TRUFAS}). The bottom panels shows the ratio between periods. The circles are the clear detections where the difference is less than 5%. A cross with a diamond marks an harmonic or subharmonic (defined as in the BLS case). The condition for harmonic or subharmonic was relaxed to $R < 0.05$.

The number of detections up to the third most significant peak of the power spectrum (taking into account harmonics and subharmonics) was 90 candidates (52.0 %) and the number of non-detections was 83 candidates (48 %). The OGLE planets OGLE-TR-10, 111, and 113 were found. The total execution time was 34 min.

We ran another test, but this time we fixed the scale to a corresponding transit duration of 2.0 hours and we raised the threshold to 3.0 times the RMS of the power spectrum. The total number of candidates found now was 88 (51%) and the number of failures was 85 (49%). Again, the OGLE planets OGLE-TR-10, 111, and 113 were among the detections. The execution time was 13.00 min.

We ran another test in which we just changed the threshold over the RMS to a value of 10. We found out that this higher threshold lost almost all candidates, and, in several cases there is not even one positive detection in the power spectrum.

In conclusion, for these type of OGLE ground-based observations, fixing the scale to an expected reasonable value (2.0 hours in this case) gives good results. Moreover, the execution time decreases almost 3 times. To raise the threshold to 3.0 times the RMS of the power spectrum helps to detect the real period and not harmonics or subharmonics. We restrict our results in the following to these parameters.

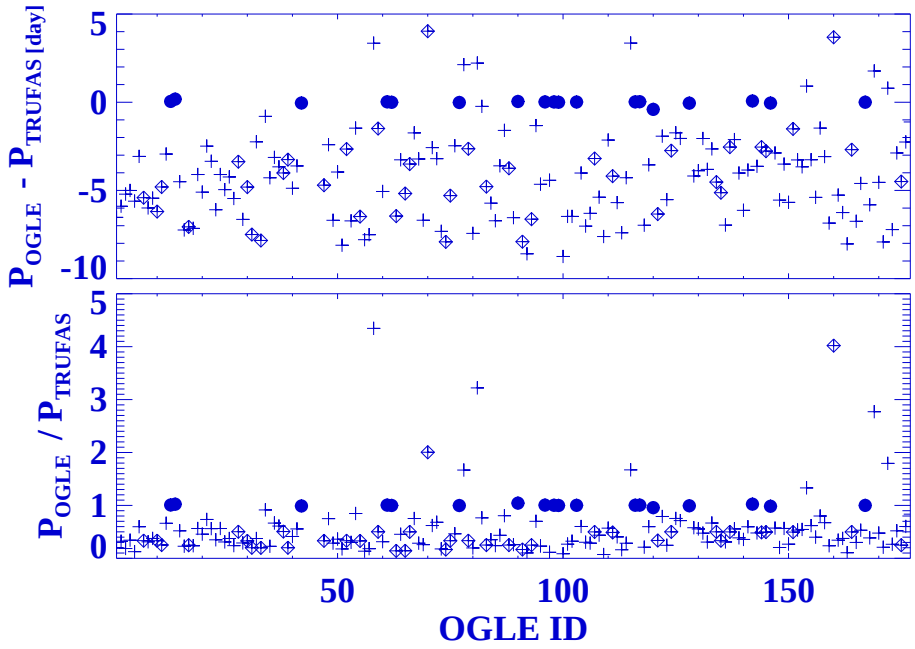

Figure 5.16: Comparison of the period determined by the OGLE team and our results using TRUFAS. In the top panel the difference is plotted. The circles are clear detections (difference is less than 5%). A cross with a diamond marks an harmonic or subharmonic. A cross (in the upper panel) shows a non-detection of the first peak of the spectrum. The bottom panel shows the ratio of period determined by OGLE and our results.

5.3.2.1 Detection efficiency

TRUFAS is tailored for continuous observations, so the introduction of gaps in OGLE-type ground-based observations is clearly a disadvantage. It is possible anyway to study its efficiency as a function of some light curve parameters that could indicate why TRUFAS is not working.

Figure 5.17 shows the TRUFAS detection efficiency as a function of number of points in transit. No clear conclusion can be drawn from it, since the efficiency changes rapidly with the number of points. However, despite the scatter, there is a correlation with the number of points in transit.

Figure 5.18 shows the TRUFAS detection efficiency as a function of number of transits present in the light curve. There is no clear dependence between TRUFAS efficiency and number of transits for OGLE light curves. It is approximately constant with an average value of 57%.

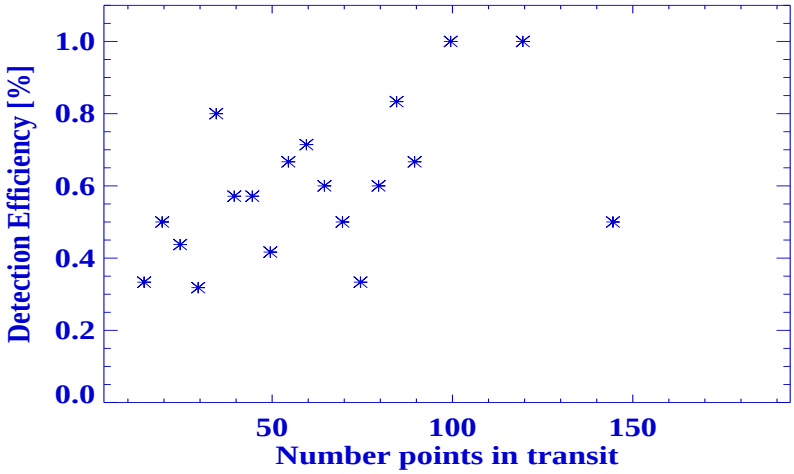

Figure 5.17: TRUFAS detection efficiency as a function of number of points in transit.

In conclusion, no correlation is seen between number of transits and TRUFAS efficiency for OGLE-type ground-based observations.

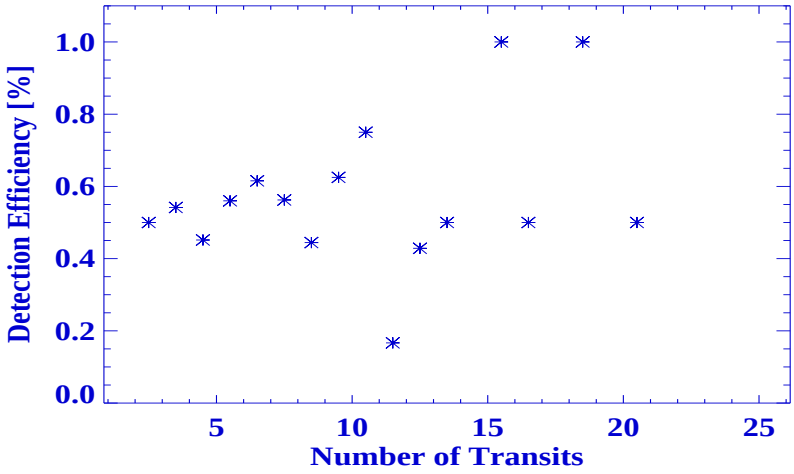

Figure 5.18: TRUFAS detection efficiency as a function of number of transits.

Figure 5.19 shows the TRUFAS detection efficiency as a function of transit depth. Despite the scatter one can see a correlation with transit depth. Evidently, low transit depth challenges TRUFAS; its efficiency decreases to 13% when the transit depth is below 0.75%. Consequently,

the transit depth is a limiting factor of TRUFAS for OGLE-type ground-based observations.

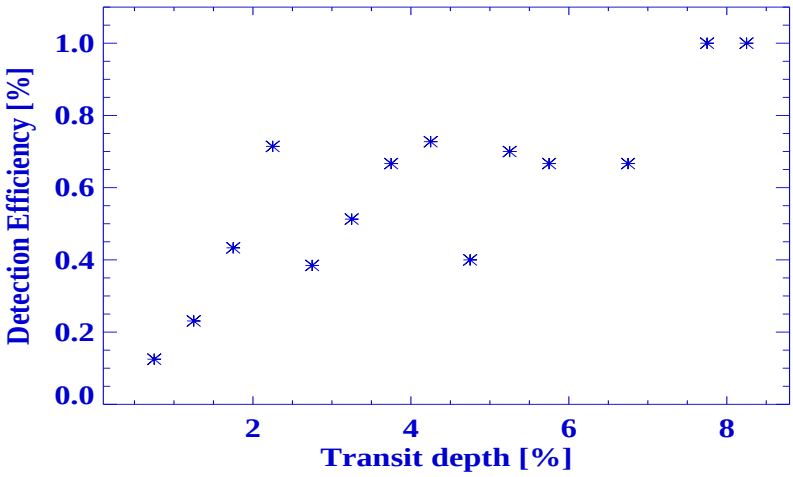

Figure 5.19: TRUFAS detection efficiency as a function of transit depth.

5.3.3 Conclusions

Using OGLE light curves we have tested the performance of the two implemented transit detection algorithms: BOX-fitting least square (BLS) algorithm (Kovács et al., 2002) and TRUFAS algorithm (Régulo et al., 2007).

The implementation of BLS works quite well. It detects 98.8% of the proposed candidates of the OGLE survey. Shallower transits and a high level of systematics challenge BLS operation.

The implementation of TRUFAS is also operating and the results are not encouraging (around 50% detection efficiency in the best run). Transits with small depth (below 0.75%) make TRUFAS detection fail.

In summary, BLS algorithm is superior to TRUFAS algorithm for these type of OGLE ground-based observations.

6

Pan-Planets simulations

The Panoramic Survey Telescope and Rapid Response System (PanSTARRS) is an Air Force funded project with the original goal of discovering and characterizing Earth-approaching objects that might pose a danger to our planet. The prototype mission PanSTARRS1 (PS1) is a 1.8m telescope at the Haleakala Observatories (Maui, Hawai). PS1 is equipped with a 7 sq.deg. CCD camera on a 1.4 Gigapixel array. PS1 is monitoring 3π rad of the sky over a 3.5 yr period.

The PS1 science consortium (USA, Germany, UK, and TAIWAN) has defined 12 Key Science Projects, one of which is the Pan-Planets survey which is led by MPIA (PIs are Afonso & Henning). It will have a total of 120h per year during the 3.5 yr lifetime of the survey. Pan-Planets will observe between 3 and 7 fields using observational blocks of 1h or 3h. More details about the Pan-Planets survey can be found in Afonso & Henning (2007).

6.1 Simulations

We have analyzed a set of first year Pan-Planets simulated light curves (Koppenhoefer et al., 2009) in order to study the efficiency of the transit detection algorithms (TDA) based on different survey strategies. We want to examine different survey parameters like number of fields to be observed (3, 5 or 7) and length of the observing block (1h or 3h).

Different number of fields correspond to different time sampling in the light curves. It is also taken into account the loss of observations due to bad weather or technical problems, as 33% of the total time. Exposure and read-out time will be 30s and 10s respectively with a target magnitude range from 13.5 to 16.5 in the Johnson V-band.

All simulated light curves contain planets of different sizes (1.0 - 1.25 R_J), transiting around main sequence stars (radii between 0.21 and 3.8 $R_\odot$, masses between 0.18 and 1.86 $M_\odot$) with

different periods (3.0 - 5.0 days). Different system configurations translate into light curves with different parameters. Therefore, it is possible to establish the efficiency of the TDA based on parameters such as number of transits, number of points in transit, and signal-to-noise ratio of the transit. Finally, correlated noise (see chapter 3) is included by adding superimposed sine waves of different wavelengths, following the Red Noise Model 4 of Koppenhoefer et al. (2009) (see section 3.4 of Koppenhoefer et al. 2009).

Table 6.1 summarizes the different sets of simulated light curves to be analyzed. For more details about the simulations, see Koppenhoefer et al. (2009).

Table 6.1: Summary of the set of simulated light curves.

ID	Fields	Block Size [h]	Time Sampling [s]	Number of points per light curve	Number of simulations[1]
1	3	1	120	2430	28878
2	3	3	120	2430	14839
3	5	1	200	1458	19309
4	5	3	200	1458	10219
5	7	1	280	1041	14585
6	7	3	280	1041	7969

[1] It corresponds to the total number of simulated light curves for that observational strategy

6.2 Box-Fitting runs

We ran BLS three times with different parameters, in order to find the most appropriate run. In all runs we searched for periods between 1.05 and 10.0 days and fractional transit lengths (transit duration divided by period) between 0.015 and 0.15. Table 6.2 shows the different combination of parameters used in the RUNs of BLS. Tables 6.3, 6.4, and 6.5 show the results of RUNs 1, 2, and 3 respectively. The percentages of detections are, in general, similar for every data set, thus no RUN is superior in this respect. The execution time of RUN-1 is almost half of the execution time of RUN-2 and RUN-3.

We have studied the additional candidates found by one RUN and not found by the other. Table 6.6 shows the comparison between BLS runs for the data set with 3 fields and 1h blocks. It shows the percentage of detections found in one given run, but not in other runs. For example, 95.8% of the detections of RUN-2 are contained in the detections of RUN-1, and 4.2% of the detections of RUN-2 are not contained in the detections of RUN-1. The latter number is the

Table 6.2: Parameters used in the RUNs of BLS.

RUN	Number of bins[1]	Number of frequencies[2]
RUN1	501	2501
RUN2	200	10000
RUN3	500	5000

1Number of bins in the phase-folded light curve.

2Number of frequencies between 1.05 [d] and 10.0 [d].

Table 6.3: BLS Results of RUN 1.

Data Set	Detection [%]	Non-Detection [%]	Execution Time [hr]
3fields & 1h	82.0	18.0	134
3fields & 3h	65.6	34.4	66
5fields & 1h	83.5	16.5	83
5fields & 3h	68.9	31.1	28
7fields & 1h	84.0	16.0	38
7fields & 3h	70.7	29.3	21

Table 6.4: BLS Results of RUN 2.

Data Set	Detection [%]	Non-Detection [%]	Execution Time [hr]
3fields & 1h	80.9	19.1	279
3fields & 3h	65.6	34.4	140
5fields & 1h	83.1	16.9	155
5fields & 3h	69.5	30.5	75
7fields & 1h	83.6	16.4	102
7fields & 3h	71.2	28.8	54

important one, because it indicates the added value of one run with respect to another. We have examined the distribution of the additional detections of one run with respect to another as a function of several parameters, such as period, transit depth, and S/N of the transit and we have found that they are distributed homogeneously in the parameter space. Figures 6.1, 6.2, and 6.3 show the additional efficiency of RUN-2 over RUN-1 as a function of period, transit depth, and S/N respectively. The small correlation with transit depth is due to the fact that bins corresponding to high transit depths have less candidates included (only 3% of the total

Table 6.5: BLS Results of RUN 3.

Data Set	Detection [%]	Non-Detection [%]	Execution Time [hr]
3fields & 1h	82.6	17.4	283
3fields & 3h	67.4	32.6	104
5fields & 1h	84.6	15.4	119
5fields & 3h	71.4	28.6	68
7fields & 1h	85.2	14.8	102
7fields & 3h	73.2	26.8	54

additional detections have transit depths bigger than 4%). The additional efficiency distributions were also homogeneous when we compared RUN-1 with RUN-3.

Table 6.6: Comparison of detections in different RUNs.

	RUN-1		RUN-2		RUN-3	
	%Contained	%Non-Contained	%Contained	%Non-Contained	%Contained	%Non-Contained
RUN-1	-	-	95.8	4.2	95.5	4.5
RUN-2	94.5	5.5	-	-	96.2	3.8
RUN-3	96.1	3.9	98.2	1.8	-	-

A final test was to check the robustness of BLS results. We ran it again with the same input parameters of RUN-1. The variation of the results was less than 0.001% in all data-sets. Thus, BLS results are numerically robust.

In conclusion, since RUN-1 was two times faster than RUN-2 and RUN-3, and the additional detections of RUN-2 and RUN-3 were homogeneously distributed along the parameter space, we prefered RUN-1 and restricted our results to it.

A candidate was defined as detected when its BLS period was within 5% of the simulated period, or when it was an harmonic or subharmonic (checked until the third peak of the BLS spectrum). Figure 6.6 shows the global BLS results for the set of light curves with 7 fields and 1h blocks. The bottom panels show the ratio between the simulated periods and the ones found by BLS and vice versa; harmonics and subharmonics are seen as peaks at integer numbers. Figures 6.4 and 6.5 show examples of Pan-Planets light curves where BLS detected and failed to find the transit respectively.

Figure 6.7 shows the Signal Detection Efficiency (SDE) as a function of S/N without (top) and with (bottom) red noise included, for the case of 7 fields and 1h blocks. The average SDE

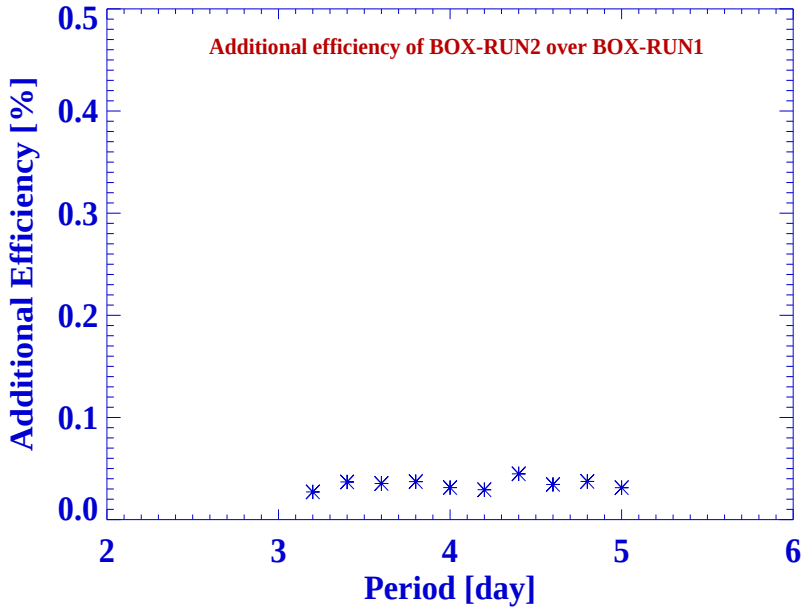

Figure 6.1: Additional efficiency of BLS-RUN2 over BLS-RUN1 as a function of period for the set with 3 fields and 1h blocks.

is *4.4* and the minimum is *1.9*. The not-so-high value of SDE is explained in part because the number of bins used in the phase-folded light curve is not so high (see Figures 3 and 4 of Kovács et al. 2002).

6.3 TRUFAS runs

We ran TRUFAS two times using different approaches. In one run we fixed the scale to an expected value, and in the other we selected the appropriate scale from a range. Fixing the scale to a reasonable value makes sense when the individual transits are not completely sampled (like Pan-Planets ground-based observations), and the automatic scale selection is therefore not as robust as desired. In both runs we searched for periods between 0.9 to 10.1 days and we looked for peaks above 3 × RMS of the power spectrum of the wavelet square transformation. The process was repeated 50 times, continuously shortening the scale to avoid detections due to random noise.

In the first run we fixed the scale to a value that corresponds to a transit duration of ∼ 2h. This transit duration is a good average if we consider the typical transiting planets that have

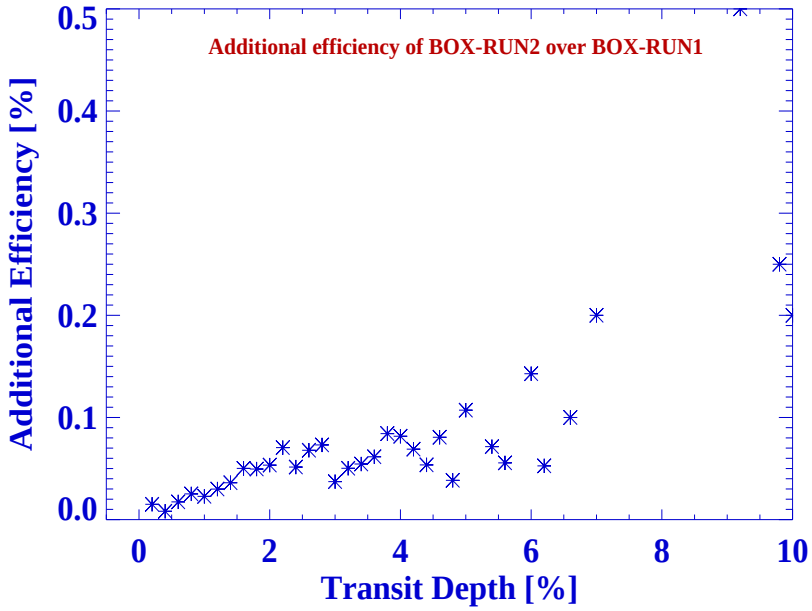

Figure 6.2: Additional efficiency of BLS-RUN2 over BLS-RUN1 as a function of transit depth for the set with 3 fields and 1h blocks.

been discovered until now (see exoplanet.eu), and the range of periods that we were examining (see Table 1 of Defaÿ et al. 2001). Table 6.7 shows the results of the first run. In all cases the efficiency is below 40%.

Table 6.7: TRUFAS Results of RUN 1.

Data Set	Detection [%]	Non-Detection [%]	Execution Time [hr]
3fields & 1h	36.8	63.2	12.0
3fields & 3h	34.1	65.9	9.0
5fields & 1h	37.1	62.9	10.0
5fields & 3h	34.8	65.2	4.8
7fields & 1h	37.0	63.0	6.4
7fields & 3h	35.0	65.0	3.8

In the second run we selected the appropriate scale from a range of expected values. Figure 6.8 shows the distribution of transit durations of the simulated set with 3 fields and 1h blocks. Transit durations range from approximately 1.0h to 4.0h ; we chose six values in this range.

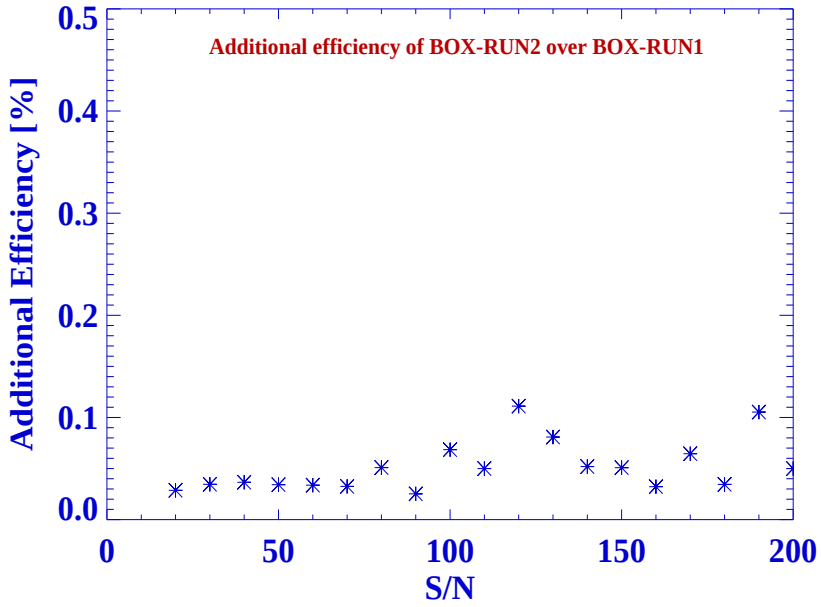

Figure 6.3: Additional efficiency of BLS-RUN2 over BLS-RUN1 as a function of S/N for the set with 3 fields and 1h blocks.

Similar distributions of transit durations are observed in the other sets of Pan-Planets simulated light curves. Table 6.8 shows the results of the second run. Again, the efficiency is below 40% in all cases.

Table 6.8: TRUFAS Results of RUN 2.

Data Set	Detection [%]	Non-Detection [%]	Execution Time [hr]
3fields & 1h	37.1	62.9	17.1
3fields & 3h	34.2	65.8	9.6
5fields & 1h	37.5	62.5	11.5
5fields & 3h	34.4	65.6	6.6
7fields & 1h	37.3	62.7	8.6
7fields & 3h	34.6	65.4	5.1

The efficiency in both cases is almost the same for all data sets. RUN-1 is faster than RUN-2. Moreover, as was mentioned previously, RUN-1 just took one default expected scale. This has the advantage that we do not need any prior knowledge of the transit duration distribution.

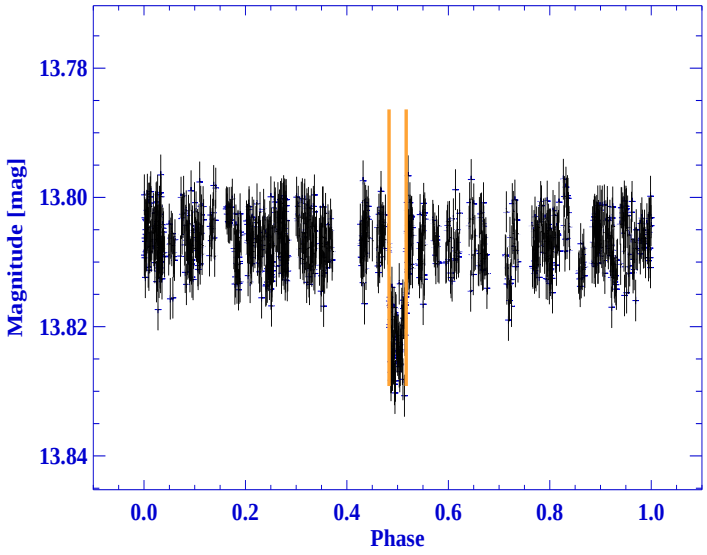

Figure 6.4: Pan-Planets light curve with a positive BLS detection. The lines represent the position of the transit.

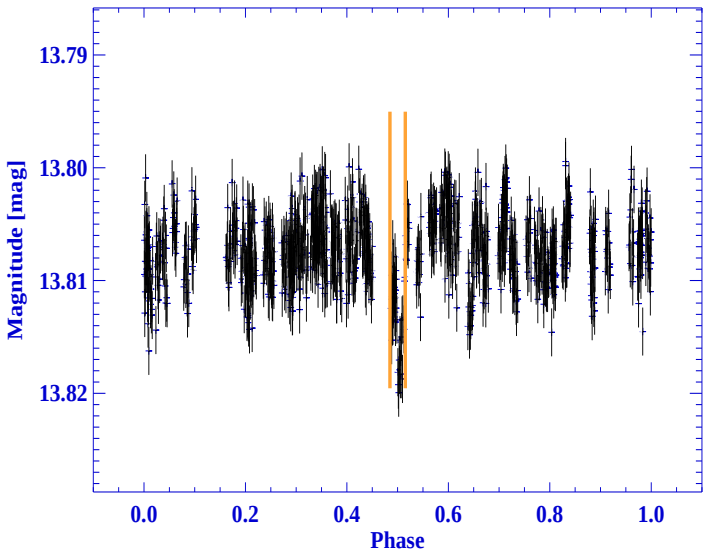

Figure 6.5: Pan-Planets light curve with a negative BLS detection. The lines represent the position of the transit.

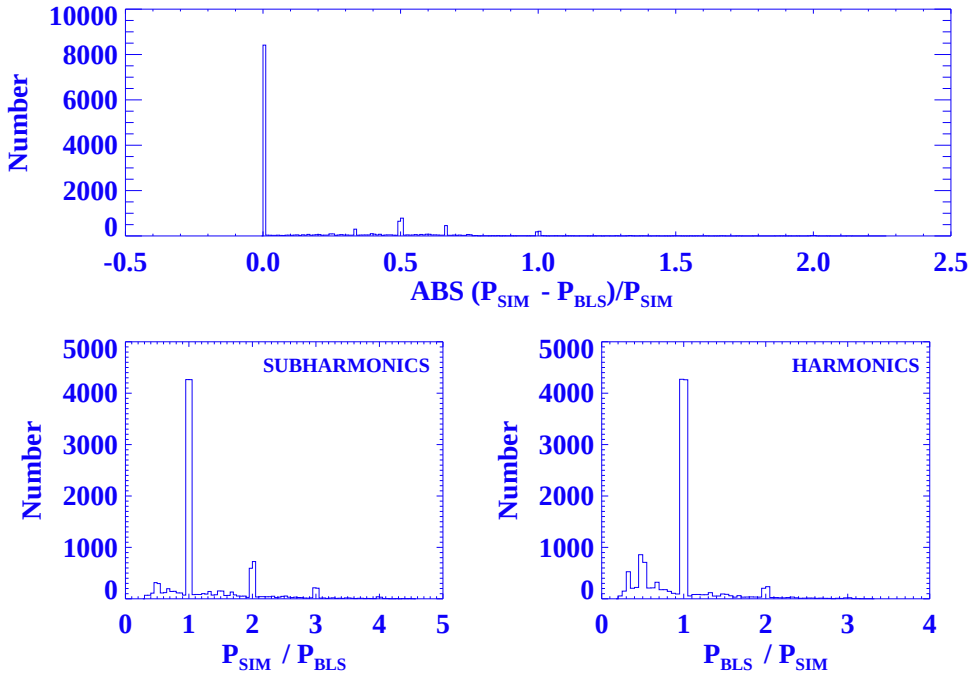

Figure 6.6: Global results of BLS for the 7 fields and 1h blocks strategy. Upper panel shows the the fractional difference between the simulated periods and the ones found by our BLS. Bottom panels show the ratio between the simulated periods and the ones found by our BLS and vice versa. Peaks at integer numbers are harmonics or subharmonics.

Like the BLS case, we studied the additional candidates found by one TRUFAS RUN and not found by the other. Table 6.9 shows the comparison between TRUFAS runs. We analyzed the distribution of additional candidates found by one RUN and not found by the other and we determined that the additional candidates were homogeneously distributed along the parameter space. Figures 6.9, 6.10, and 6.11 show the additional efficiency of RUN-2 over RUN-1 as a function of period, transit depth, and S/N respectively. It is seen that the parameters were homogeneously distributed, and no special efficiency trend was observed.

In conclusion, since both RUNs detections were equally efficient, but RUN-1 was faster and it did not assume any prior parameter distribution, we prefered RUN-1. We restricted our results to it.

Figure 6.12 shows the global results of TRUFAS for the case of 7 fields and 1h block strategy. As can be seen, most of the results were not in agreement, reaching just 37% of detections.

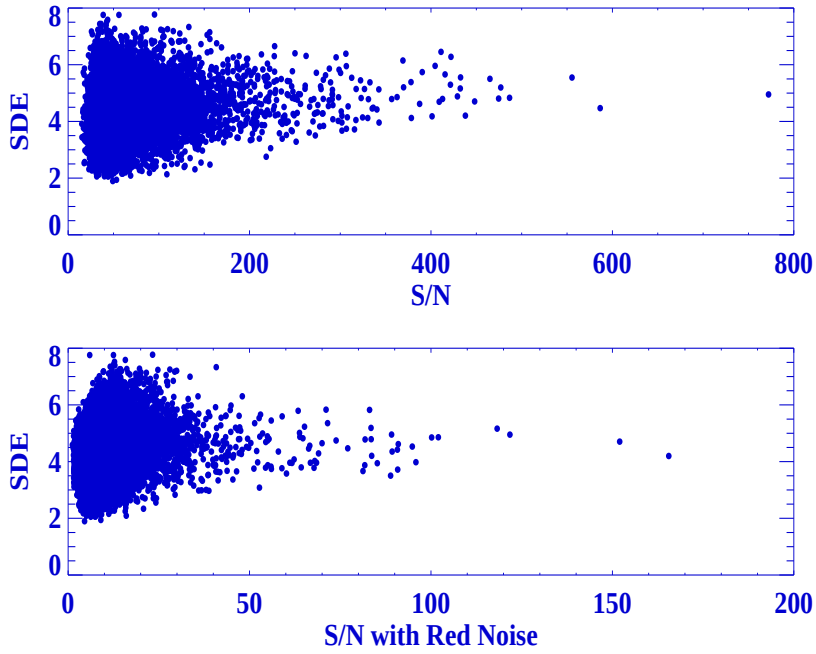

Figure 6.7: Signal detection Efficiency as a function of S/N without (top panel) and with (bottom panel) red noise for the 7 fields and 1h blocks strategy.

Table 6.9: Comparison between the two TRUFAS RUNs.

	RUN-1		RUN-2	
	%Contained	%Non-Contained	%Contained	%Non-Contained
RUN-1	-	-	90.5	9.5
RUN-2	91.1	8.9	-	-

Peaks at integer numbers are harmonics or subharmonics. Figures 6.13 and 6.14 show examples of Pan-Planets light curves where TRUFAS detected and failed to find the transit.

6.4 Efficiency of the transit detection algorithms

The Pan-Planets simulations are an ideal set to test the efficiency of our transit detection algorithms with good statistics. As in the OGLE case, we want to establish the efficiency of both algorithms as a function of several parameters, and moreover, to determine the survey strategy that maximizes the detection efficiency.

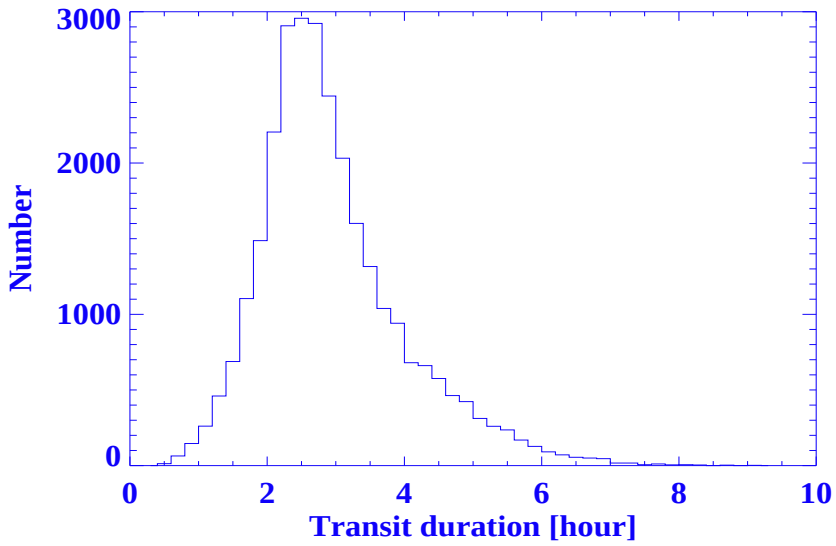

Figure 6.8: Transit duration distribution of the simulated data set with 3 fields and 1h blocks.

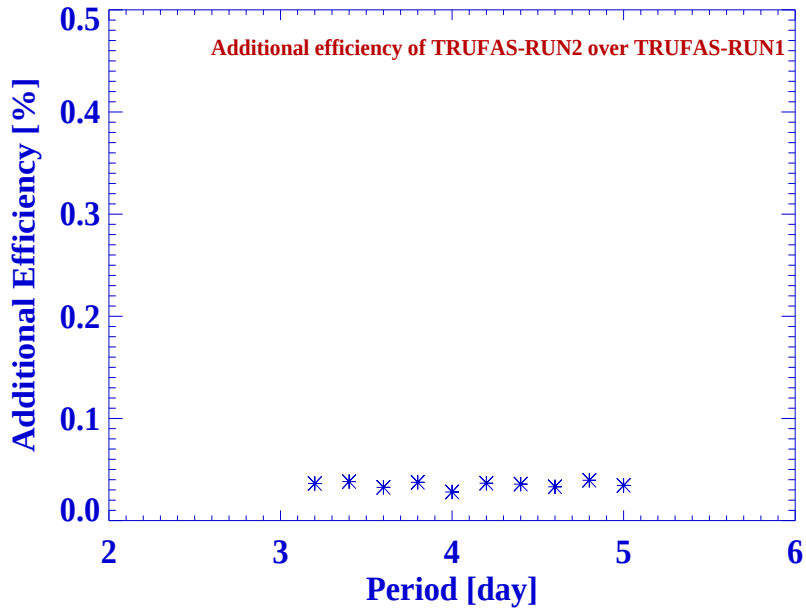

Figure 6.9: Additional efficiency of TRUFAS-RUN2 over TRUFAS-RUN1 as a function of period for the set with 3 fields 1h blocks.

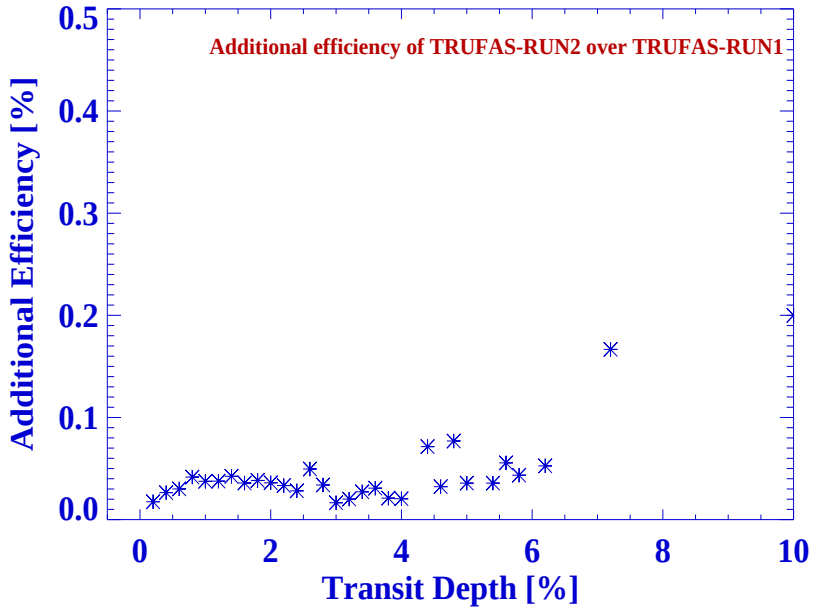

Figure 6.10: Additional efficiency of TRUFAS-RUN2 over TRUFAS-RUN1 as a function of transit depth for the set with 3 fields 1h blocks.

6.4.1 Influence of the observing block size

We have studied the influence of the size of the observing block on the detections. As the planet orbits, it spends a certain fraction of time in transit. The time spent in transit depends on the period, inclination of the system, and radius of the host star. With 1h blocks, we can sample much better the light curve (more transits). But on the other hand, we miss the complete transit (which typically lasts between 1h and 3h, Defaÿ et al. 2001), and it is only possible to observe part of it. With 3h blocks, the observational windows are less and therefore we decrease the period sampling (we observe less transits). But this has the advantage that we can observe complete transits.

In the case of BLS (see Table 6.3), it is clear that the 1h blocks are superior to 3h blocks, because more transits are observed and the BLS algorithm is more efficient. Figures 6.15 and 6.16 show the BLS efficiency as a function of period for the 1h and 3h block strategies respectively. The efficiency for the 3h block strategy is always lower than the efficiency for the 1h block strategy. Figures 6.17 and 6.18 show the BLS efficiency as a function of number of transits for the 1h and 3h block strategies respectively. As was expected, the number of observed transits

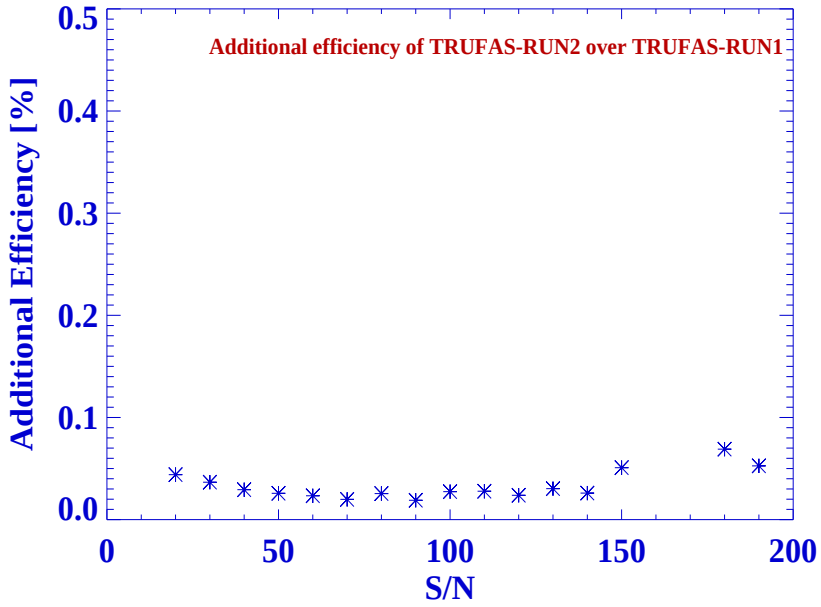

Figure 6.11: Additional efficiency of TRUFAS-RUN2 over TRUFAS-RUN1 as a function of S/N for the set with 3 fields 1h blocks.

is bigger for 1h blocks. The efficiency is in general homogeneous for both cases, with a slight increase around 10 transits, where it starts to rise. The average efficiency is 89.1% for the 1h blocks case and 77.8% for the 3h blocks case. In conclusion, the 1h blocks strategy allows to observe more transits, and therefore, the BLS algorithm works better and its average detection efficiency is bigger.

In Table 6.7 we can see that the TRUFAS detection efficiency of both strategies is very similar; the difference is just around 2% in all cases. Figures 6.19 and 6.20 show the TRUFAS efficiency as a function of period for the 1h and 3h block strategies respectively. The efficiency is homogeneous in both cases, with averages of 37.5% and 35.1% for the 1h blocks and 3h blocks strategies respectively. Figures 6.21 and 6.22 show the TRUFAS detection efficiency as a function of number of transits for the 1h and 3h block strategies respectively. As can be seen, TRUFAS is more efficient if we observe more transits. This is because it operates based on the presence of individual transits and not on the phase-folded light curve. An individual transit is represented as one peak in the wavelet transformation (WT). Consequently, more transits mean more peaks in the WT, and as a result, it is easier to find the period of the planet in the power

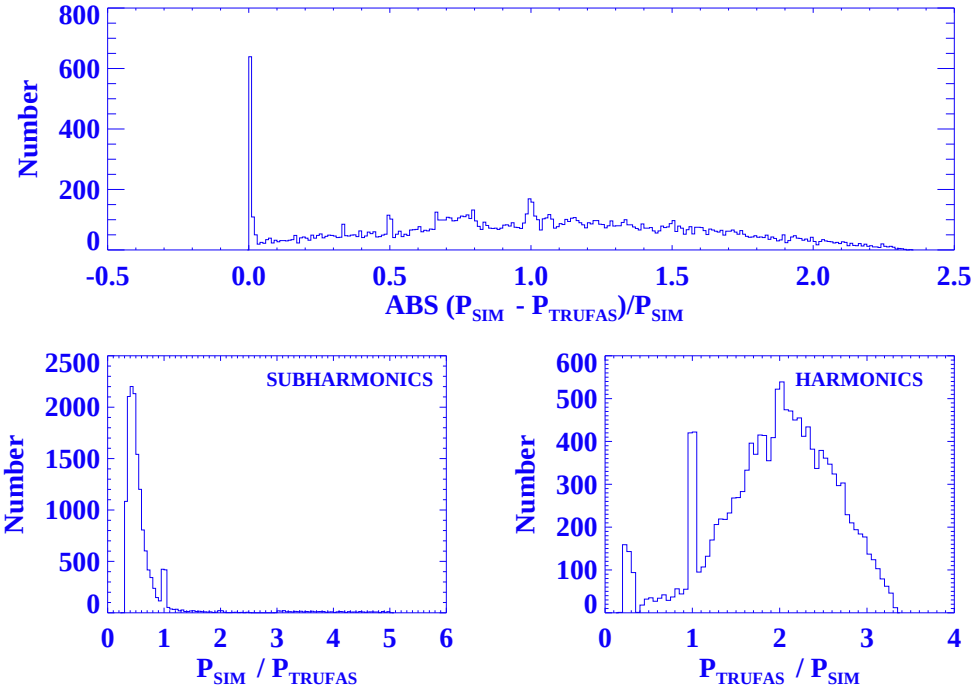

Figure 6.12: Global results of TRUFAS for the 7 fields and 1h blocks strategy. Upper panel shows the the fractional difference between the simulated periods and the ones found by our TRUFAS. Bottom panels show the ratio between the simulated periods and the ones found by our TRUFAS and vice versa. There was not much agreement between the simulated results and TRUFAS results.

spectrum of this WT.

In summary, both algorithms work better with 1h blocks. In both cases, more transits make the algorithms more efficient. We restrict our results to the 1h blocks strategy in the following.

6.4.2 Influence of time sampling

When more fields are observed (given a fixed amount of observing time), the time sampling is degraded. Thus we have less points per light curve, and in consequence, the S/N of the transit is decreased. On the other hand, observing more fields allows to monitor more stars, which permits the detection of more transits. Tables 6.3 and 6.7 show that the percentage of detected planets of BLS and TRUFAS is almost the same for the three strategies (3, 5, and 7 fields with 1h blocks).

Figures 6.23 and 6.24 show the BLS and TRUFAS detection efficiencies as a function of

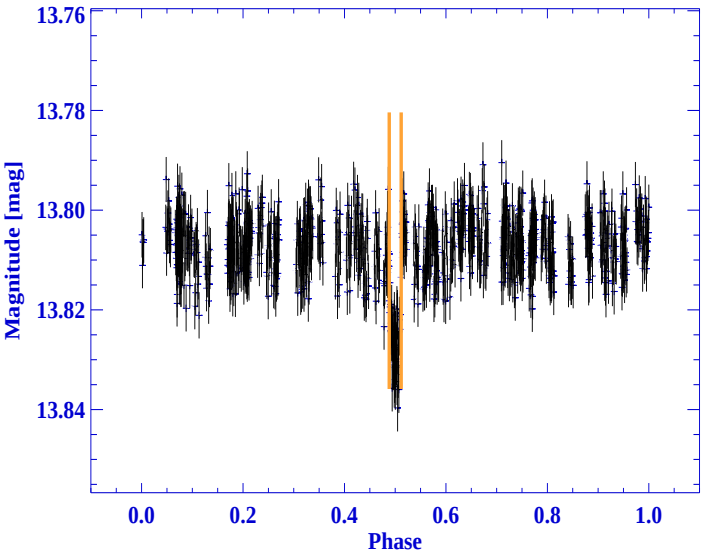

Figure 6.13: Pan-Planets light curve with a positive TRUFAS detection. The vertical lines represent the position of the transit.

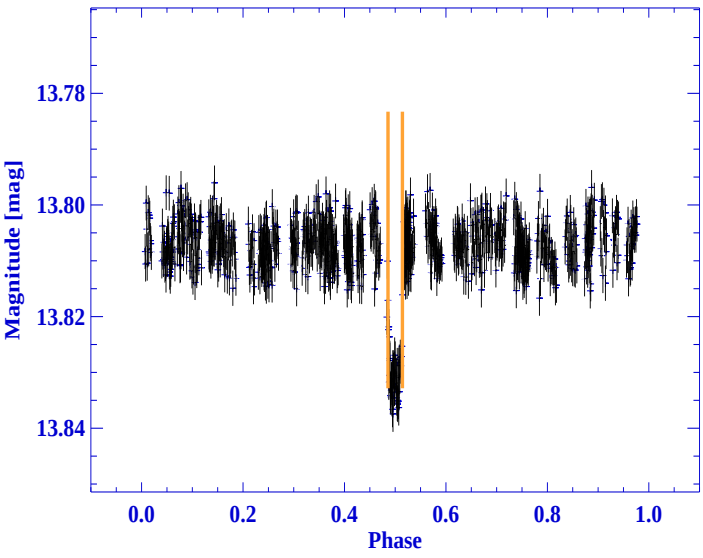

Figure 6.14: Pan-Planets light curve with a negative TRUFAS detection. The vertical lines represent the position of the transit.

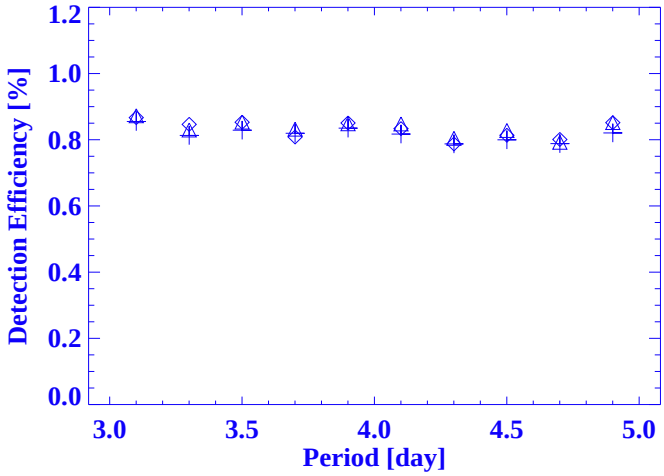

Figure 6.15: BLS efficiency as a function of period for the 1h block strategies. Crosses, diamonds, and triangles represent the 3, 5, and 7 field strategies respectively. The average detection efficiency is 82.8%.

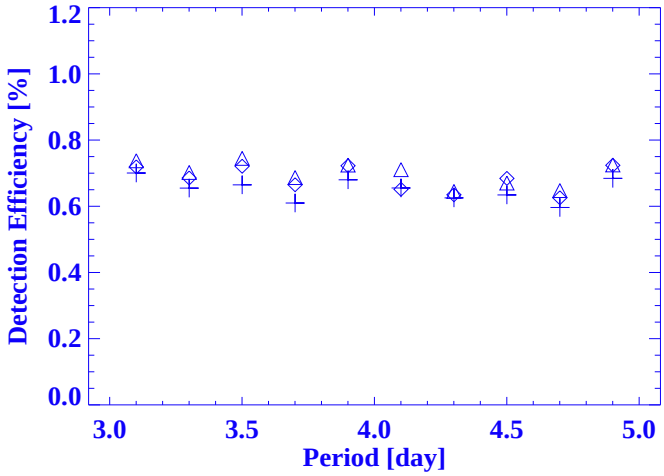

Figure 6.16: BLS efficiency as a function of period for the 3h block strategies. Crosses, diamonds, and triangles represent the 3, 5, and 7 field strategies respectively. The average detection efficiency is 67.8%

number of points in transit for the 1h blocks strategies. In the BLS case, the efficiency is homogeneous with average detection efficiencies of 87.9%, 88.0%, and 89.2% for the strategies with 3, 5, and 7 fields respectively. Although there is a small efficiency decrease with less points

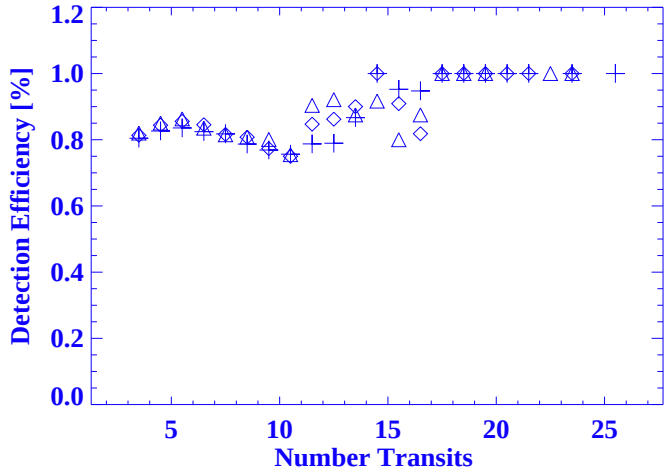

Figure 6.17: BLS efficiency as a function of number of transits for the 1h block strategies. Crosses, diamonds, and triangles represent the 3, 5, and 7 field strategies respectively. The average detection efficiency is 89.1%.

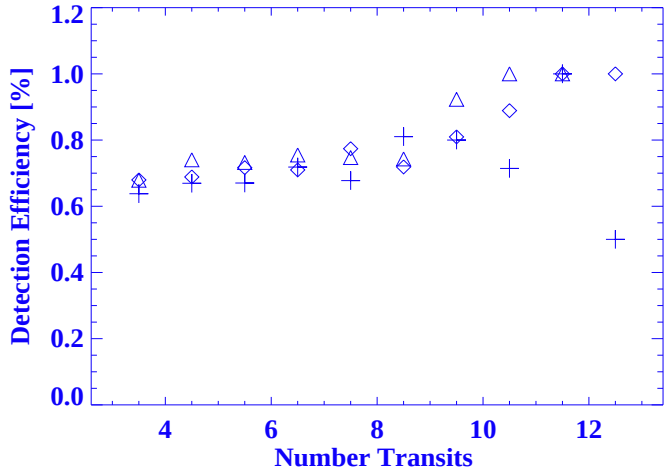

Figure 6.18: BLS efficiency as a function of number of transits for the 3h block strategies. Crosses, diamonds, and triangles represent the 3, 5, and 7 field strategies respectively. The average detection efficiency is 77.8%.

in transit, there is no clear indication that it is a limiting factor of BLS, since this trend is observed in all cases. In the TRUFAS case, the efficiency increases when we have more points in transit. The average detection efficiencies are 67.2%, 63.5%, and 69.1% for the strategies with

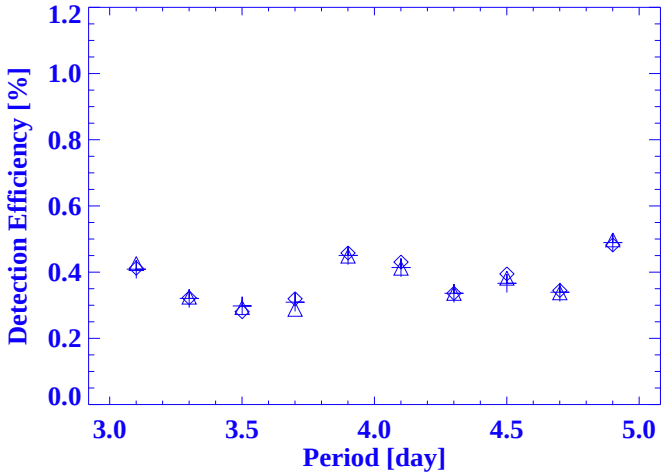

Figure 6.19: TRUFAS efficiency as a function of period for the 1h block strategies. Crosses, diamonds, and triangles represent the 3, 5, and 7 field strategies respectively. The average detection efficiency is 37.5%.

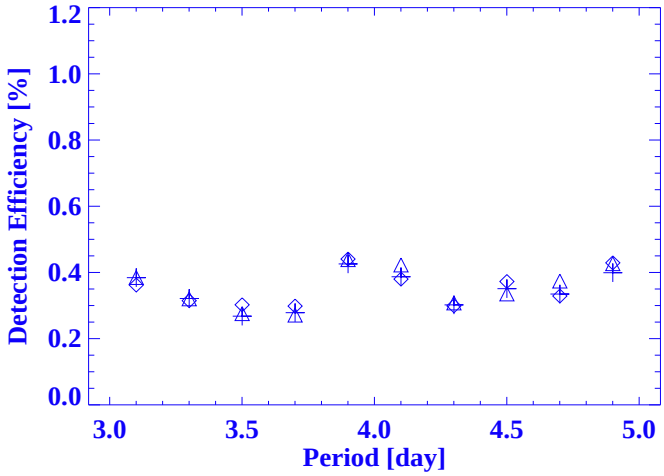

Figure 6.20: TRUFAS efficiency as a function of period for the 3h block strategies. Crosses, diamonds, and triangles represent the 3, 5, and 7 field strategies respectively. The average detection efficiency is 35.1%

3, 5, and 7 fields respectively. There is no obvious evidence that one strategy is better than the other for TRUFAS, because in all of them the same pattern is repeated.

Figures 6.25 and 6.26 show the BLS and TRUFAS detection efficiencies as a function of S/N

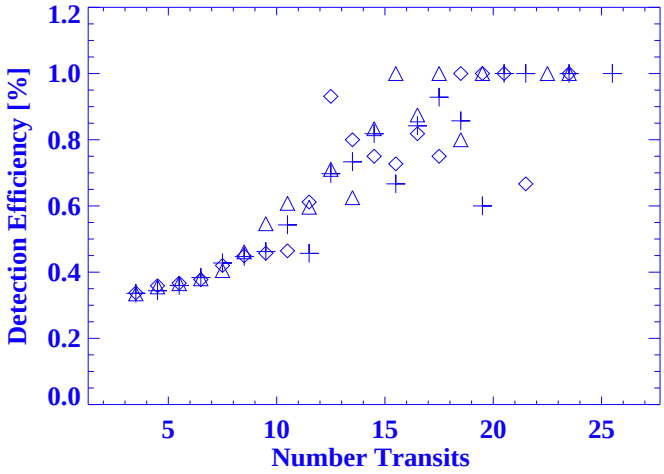

Figure 6.21: TRUFAS efficiency as a function of number of transits for the 1h block strategies. Crosses, diamonds, and triangles represent the 3, 5, and 7 field strategies respectively. The average detection efficiency is 66.8%.

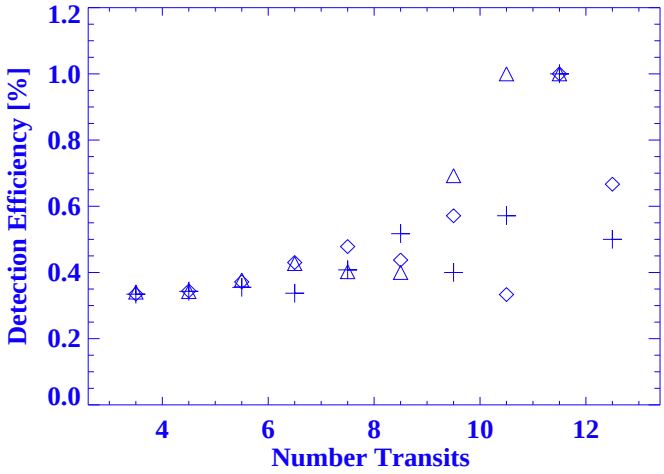

Figure 6.22: TRUFAS efficiency as a function of number of transits for the 3h block strategies. Crosses, diamonds, and triangles represent the 3, 5, and 7 field strategies respectively. The average detection efficiency is 50.9%.

of the transit for the 1h blocks strategies. In the BLS case, the efficiency is almost homogeneous and greater than 80% in all strategies, with a decrease towards small S/N (20%). In the TRUFAS case, there is no obvious trend and just scatter is observed.

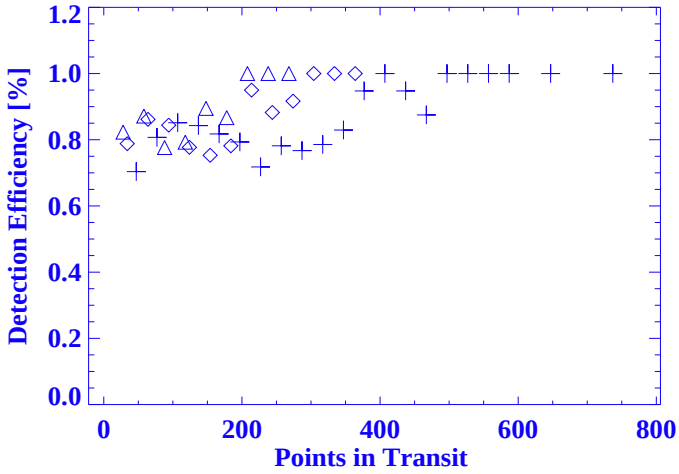

Figure 6.23: BLS efficiency as a function of number of points in transit for the 1h block strategies. Crosses, diamonds, and triangles represent the 3, 5, and 7 field strategies respectively.

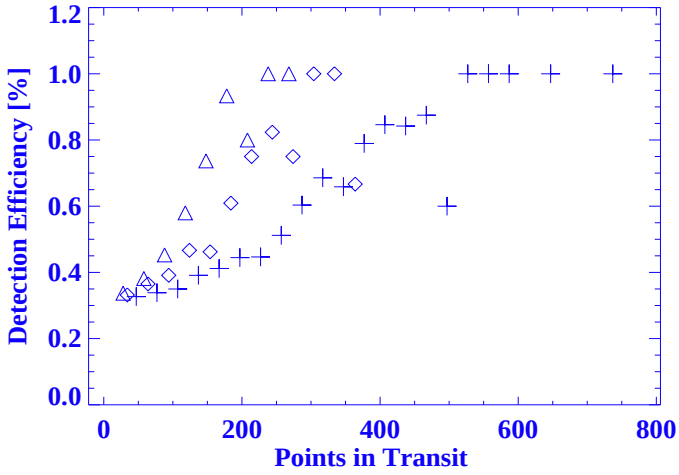

Figure 6.24: TRUFAS efficiency as a function of number of points in transit for the 1h block strategies. Crosses, diamonds, and triangles represent the 3, 5, and 7 field strategies respectively.

In summary, observing more fields is better because there are still enough points in transit, and in consequence, a good S/N can still be achieved. Moreover, the loss of time sampling is compensated by the larger number of stars that are monitored, which in turn allows the detection of more transiting planets. We will restrict our discussion to the 7 field case in the

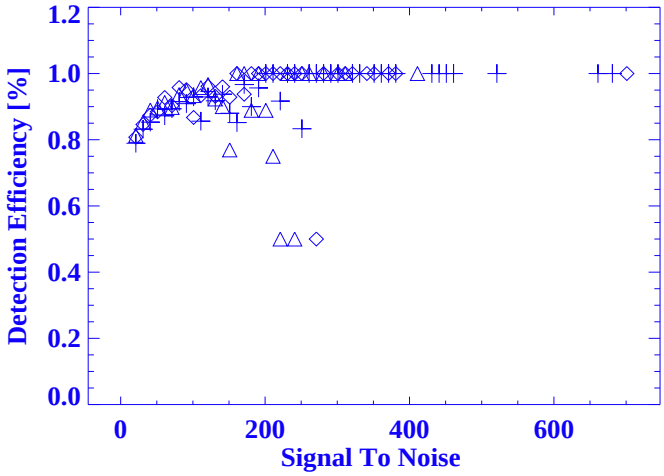

Figure 6.25: BLS efficiency as a function of S/N for the 1h block strategies. Crosses, diamonds, and triangles represent the 3, 5, and 7 field strategies respectively.

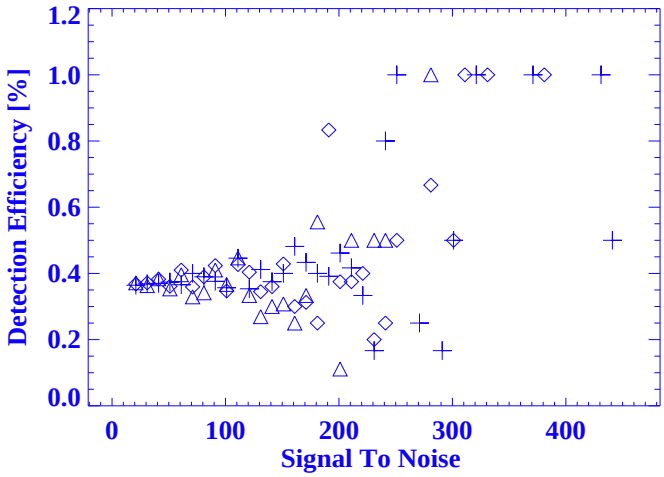

Figure 6.26: TRUFAS efficiency as a function of S/N for the 1h block strategies. Crosses, diamonds, and triangles represent the 3, 5, and 7 field strategies respectively.

following.

6.4.3 Influence of the transit depth

The transit depth of a planet is approximately given by the square ratio of planetary to stellar radius (if limb darkening is ignored). Consequently, one given planet can produce different transit depths depending on the star it is orbiting. As example, a Jupiter-like planet orbiting a Sun-like star produces a transit depth of 1%.

The Pan-Planets survey is expected to find up to 10 Hot Jupiters (periods between 3 and 5 days, radii of 1.0 - 1.25 R_J) during the first year of the survey (Koppenhoefer et al., 2009). Therefore, measuring the efficiency of the algorithms in the expected range of transit depths is very important.

Figure 6.27 shows the BLS efficiency as a function of transit depth for the 7 fields and 1h blocks strategy. It is clear that the transit depth influences the detection, because the efficiency drops very rapidly for lower transit depths. If the transit depth is below 0.3% the detection efficiency is below 35.3%.

Figure 6.28 shows the TRUFAS efficiency as a function of transit depth for the 7 fields and 1h blocks strategy. It is homogeneous along all transit depths, with an average detection efficiency of 46%. No clear dependence of the TRUFAS efficiency with transit depth is observed.

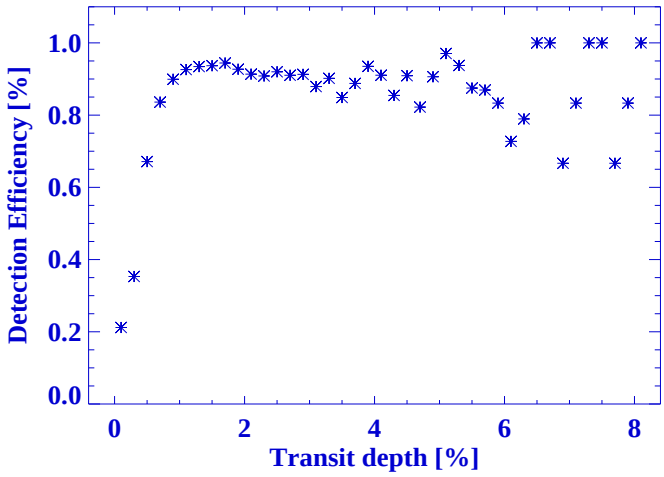

Figure 6.27: BLS detection efficiency as a function of transit depth for the 7 fields and 1h blocks strategy.

In summary, the transit depth is a very determinant factor on the detections for the case of BLS, even though it is possible to detect transits with small transit depth. TRUFAS does not

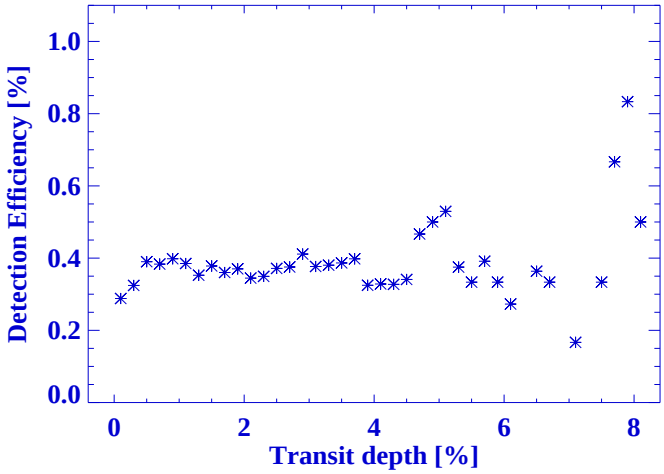

Figure 6.28: TRUFAS detection efficiency as a function of transit depth for the 7 fields and 1h blocks strategy.

show clear dependence on it, as its efficiency is fairly constant along the studied transit depths.

6.4.4 Influence of Red Noise

The correlated noise (see chapter 3) is always present at a certain level in photometric time-series (Pont et al., 2006). The correlated noise lowers the significance of a transit, because trends due to red noise resemble transit features. This effect is clearly seen in Figure 6.29, which shows the distribution of S/N taking and not taking into account the red noise for the 7 fields 1h blocks strategy (using equations 4.9 and 3.6). The minimum S/N in the simulated light curves is 16.0, a value that decreases to 1.2 if we include red noise. The average S/N in the case of 7 fields and 1h blocks is 30.0. If we consider red noise this average decreases to 11.3. In conclusion, it is clear that red noise lowers the significance of a transit.

Figure 6.30 shows the BLS detection efficiency as a function of S/N taking into account red noise. It is clearly seen that the S/N with red noise is a good indicator of the BLS efficiency and therefore, the red noise is an important limitation to be considered. If the S/N with red noise is below 5.0, the detection efficiency is below 68%.

Figure 6.31 shows the TRUFAS detection efficiency as a function of the S/N taking into account red noise. The efficiency is uniform along all S/N with the exception of the low S/N region, where a small decrease is observed. The average efficiency is 44%, and if the S/N is

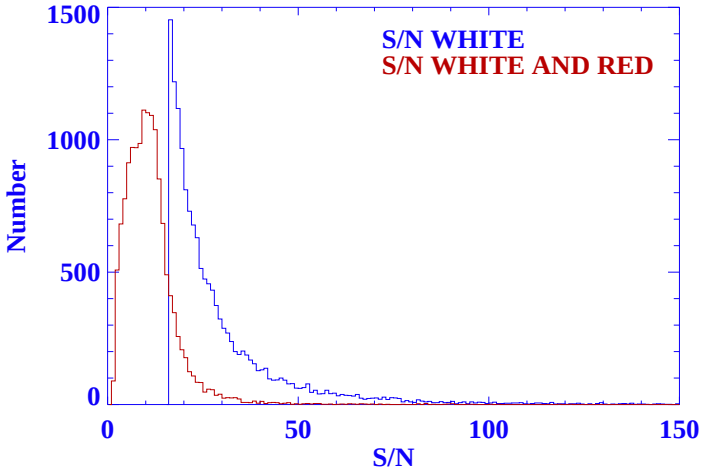

Figure 6.29: Distribution of S/N taking and not taking into account correlated noise for the 7 fields 1h blocks strategy.

below 5.0, the detection efficiency is below 37%.

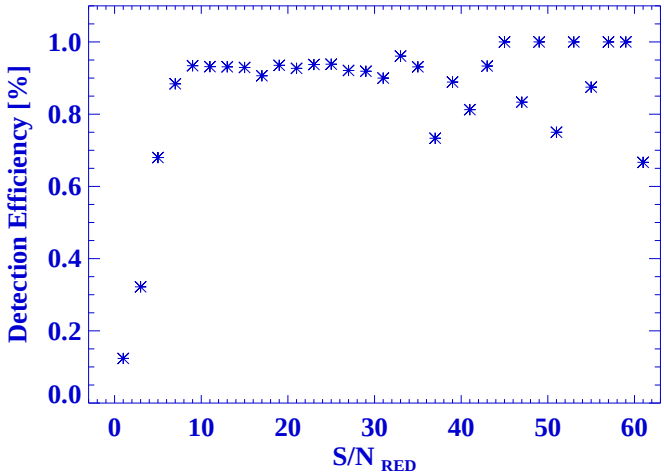

Figure 6.30: BLS efficiency as a function of S/N taking into account correlated noise for the 7 fields 1h blocks strategy.

In summary, the S/N that includes red noise is a good indicator of the detection efficiency of BLS. On the other hand, TRUFAS efficiency does not show apparent variations with it.

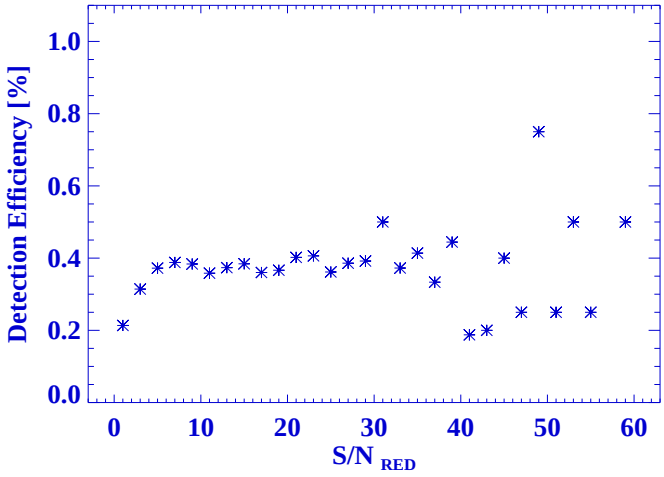

Figure 6.31: TRUFAS efficiency as a function of S/N taking into account correlated noise for the 7 fields 1h blocks strategy.

6.4.5 Conclusions

We have analyzed the performance of the two implemented transit detection algorithms (BLS and TRUFAS) using Pan-Planets simulations corresponding to the first year of the survey.

We have found that 1h blocks are better than 3h blocks for both algorithms, because more transits are observed, and the algorithms operate better when more transits are observed.

We have established that observing more fields is better, even though the time sampling is degraded. This is because the S/N is still high enough to confirm transits, while more stars are monitored.

We have evaluated the limitations of the algorithms in the selected observational strategy (7 fields and 1h blocks). We have determined that the main indicators of BLS efficiency are the transit depth and S/N with red noise. Shallower transits (low transit depth) with high level of correlated noise (red noise) challenge BLS. TRUFAS operates better when more transits are observed (which is the case of 1h blocks). Its efficiency does not change much with transit depth or S/N with red noise included.

Clearly, BOX-fitting least square algorithm is superior to TRUFAS algorithm for Pan-Planets-like ground-based observations.

6.5 Transit detection algorithms comparison

We have established that BLS is superior to TRUFAS for Pan-Planets-like observations, but TRUFAS efficiency is still significant. We want to study the properties of TRUFAS detections with respect to BLS detections in order to decide if it is worth using it.

Of the total of 5389 (37%) TRUFAS detections, 87% are contained in the 12244 (84%) BLS detections for the 7 fields and 1h blocks strategy. There are 702 TRUFAS detections (13%) that are additional. Figure 6.32 shows the additional TRUFAS detection efficiency as a function of transit depth for the 7 fields and 1h blocks strategy. As can be seen, the additional efficiency has a small increase towards lower transit depth. Its value is 24% when the transit depth is 0.2%. It is known that a transit with low depth has more chances to be due to a planet. Consequently, this additional efficiency is an argument in favor of TRUFAS.

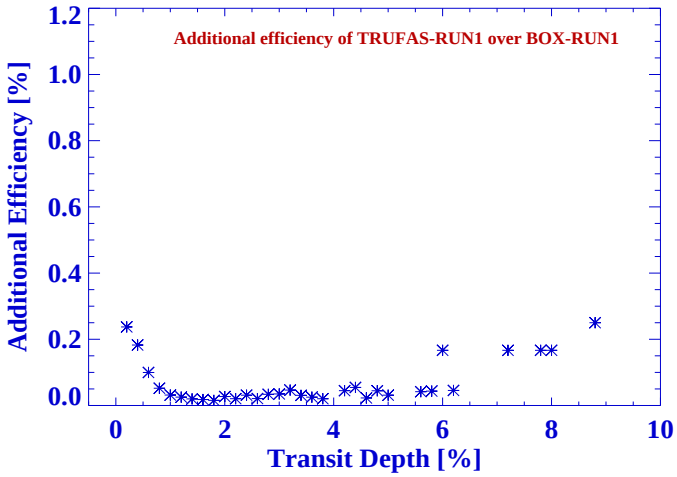

Figure 6.32: Additional detection efficiency of TRUFAS for the 7 fields and 1h blocks strategy. The additional detection efficiency of TRUFAS is 24% when the transit depth is 0.2%

Figures 6.33 and 6.34 show two examples of light curves (IDS 5929 and 8190) with low transit depth (0.2%) that were not detected by BLS but were detected by TRUFAS. Of the additional detected candidates by TRUFAS, there are 28 candidates with transit depths lower than 0.2%. In conclusion, TRUFAS could be used to detect the planets that produce a small transit depth, although the efficiency is in general low.

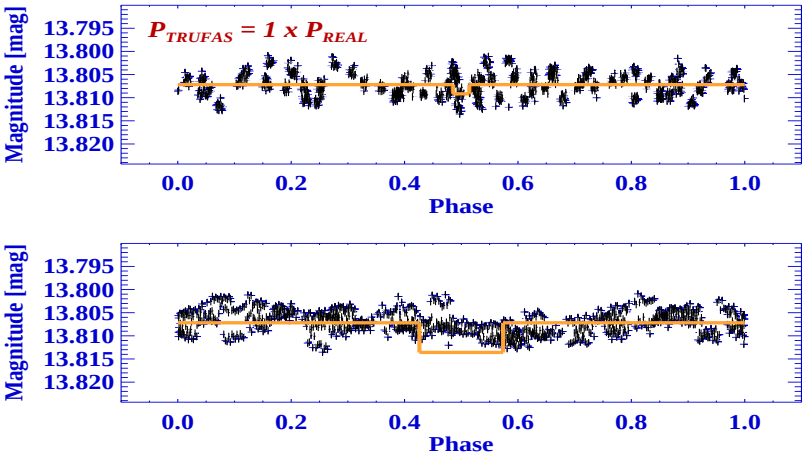

Figure 6.33: Light curve ID 5929 of the set of simulated light curves with 7 fields and 1h blocks, light curve with a transit depth of 0.2%. Top panel shows the real phase-folded light curve, The relation between the period found by TRUFAS and the real period is indicated in red. Bottom panel shows the BLS phase-folded light curve (with the wrong parameters).

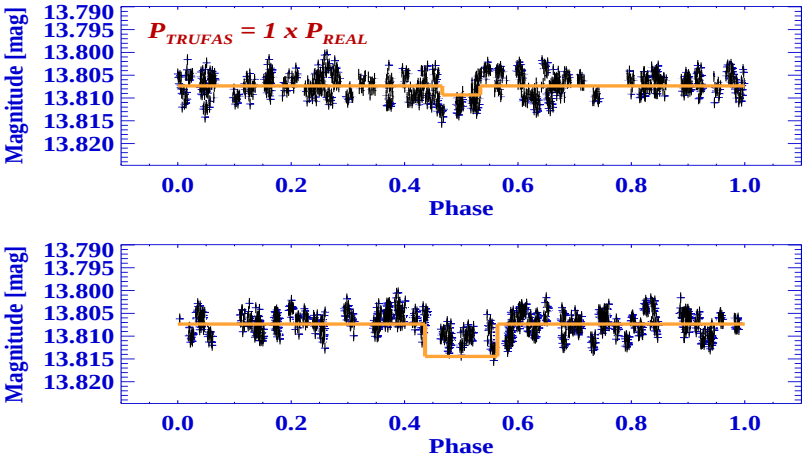

Figure 6.34: Light curve ID 8190 of the set of simulated light curves with 7 fields and 1h blocks, light curve with a transit depth of 0.2%. Top panel shows the real phase-folded light curve, the relation between the period found by TRUFAS and the real period is indicated in red. Bottom panel shows the BLS phase folded light curve (with the wrong parameters) with the BLS fit superimposed.

7

Transiting planets in LAIWO data

7.1 The LAIWO Project

LAIWO (Large Area Imager at the Wise Observatory, Baumeister et al. 2006) is a new wide-field CCD camera for the 1m Ritchey-Chretien reflector telescope at Wise Observatory in the Negev desert, Israel (see figure 7.1). LAIWO was built at the Max Planck Institute for Astronomy in Heidelberg, Germany. In terms of telescope size and field of view, the LAIWO project is similar to the OGLE survey (see chapter 5). The scientific aim of the instrument is to detect Jupiter-like extra-solar planets with the transit method (with photometry precision better than $\sim 1.5\%$ down to R=16.5 mag).

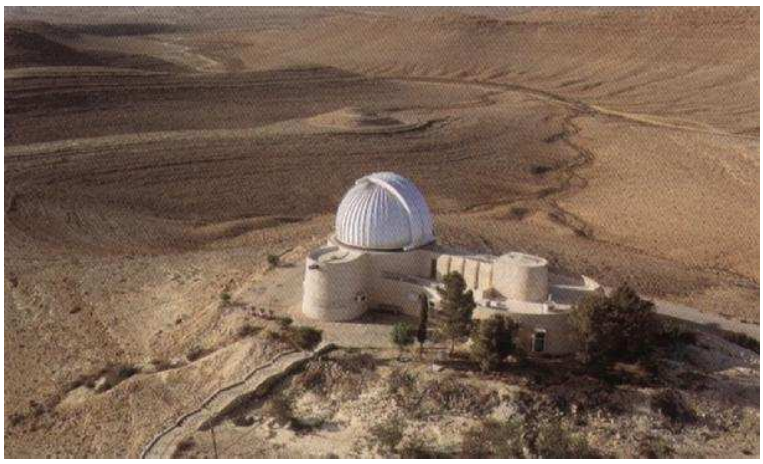

Figure 7.1: Wise Observatory in the Negev desert, Israel.

7.1.1 LAIWO camera

LAIWO is an array of four non-contiguous frontside-illuminated science CCDs with 4K × 4K pixels each. The pixel size is 15 μm and the total field of view is $59' \times 59'$. In the middle of the four science CCDs there is a 1k × 1k backside-illuminated CCD for telescope guiding. At the f/7 focus the pixel size is $0.44''$ and each CCD images a 29.5×29.5 $arcmin^2$ field. Each science CCD is connected to four output channels to reduce the read-out time, thus, each CCD is organized in four 2K × 2K quadrants which can be read-out simultaneously and individually. Figure 7.2 shows the CCD layout of the 16 quadrants in the North-East orientation.

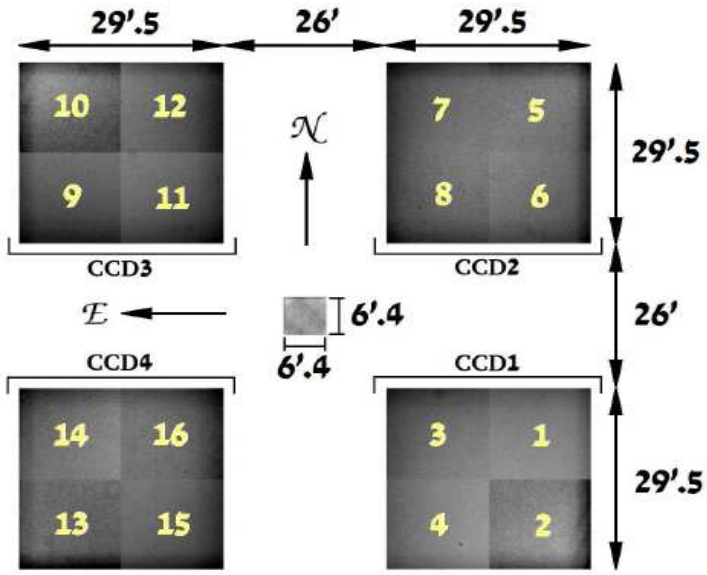

Figure 7.2: CCD Layout of the 16 LAIWO quadrants in the North-East orientation (Gorbikov et al., 2010). The relative sizes and distances are preserved.

Table 7.1 summarizes important properties of the LAIWO camera.

LAIWO images are mosaic FITS files with 16 extensions. The science CCD images are binned 2×2, thus, the quadrant size is $1k \times 1k$ and the binned pixel size is $0.86''$. The read-out time of the entire mosaic in 2×2 binning is $\sim$ 28s.

Table 7.1: Main Properties of LAIWO camera

Parameter	Value
Pixel size	$15\mu m$
FOV	$59' \times 59'$
Pixel scale	$0.44''$ per pixel
Read-out noise	$9 < RON < 19e^-$
Gain	$5e^- ADU^{-1}$

7.2 Observations

7.2.1 LAIWO field

The studied field "LAIWOVI" is located within the Cygnus-Lyra region. It is within the Kepler field (Borucki et al., 2009), making it extremely useful to add planet candidates to the Kepler target star catalog[1]. Table 7.2 summarizes some of its properties and observations.

Table 7.2: LAIWOVI properties

PARAMETER	VALUE
RA (J2000)	$19^h28^m59^s$
DEC (J2000)	$47^d58'10''$
Number of points	1046
Time sampling	3.5 [min]
Median Number of stars	13501
Median Seeing	3.2 [arcsec]

The 2009 campaign collected 1046 R-band images within a time span of $\sim$ 2 months. The integration time was 180.0s, thus the time sampling was 3.5 min. Figures 7.3, 7.4 and, 7.5 show the initial seeing, number of stars and, background distributions of the LAIWOVI field. Since the distributions are not quite Gaussian, it was decided to do a preliminary cut based on the normal mean and standard deviation (σ) statistics rather than the Gaussian statistics. Images with seeing conditions greater than $1.6 \times \sigma$ from the median value were cut. Images with number of stars smaller than $1.6 \times \sigma$ from the median value were also cut. No cut in background conditions was performed at this stage. Figure 7.6 show the correlation between number of stars and seeing. A slight correlation is seen, which is consistent with the fact that

[1]See http://nsted.ipac.caltech.edu/data/NStED/kic_columns.html

more stars will be detected with better seeing conditions. It is also important to notice that different quadrants exhibit different properties, although some of them belong to the same CCD. Thus, it is important to analyze every quadrant individually.

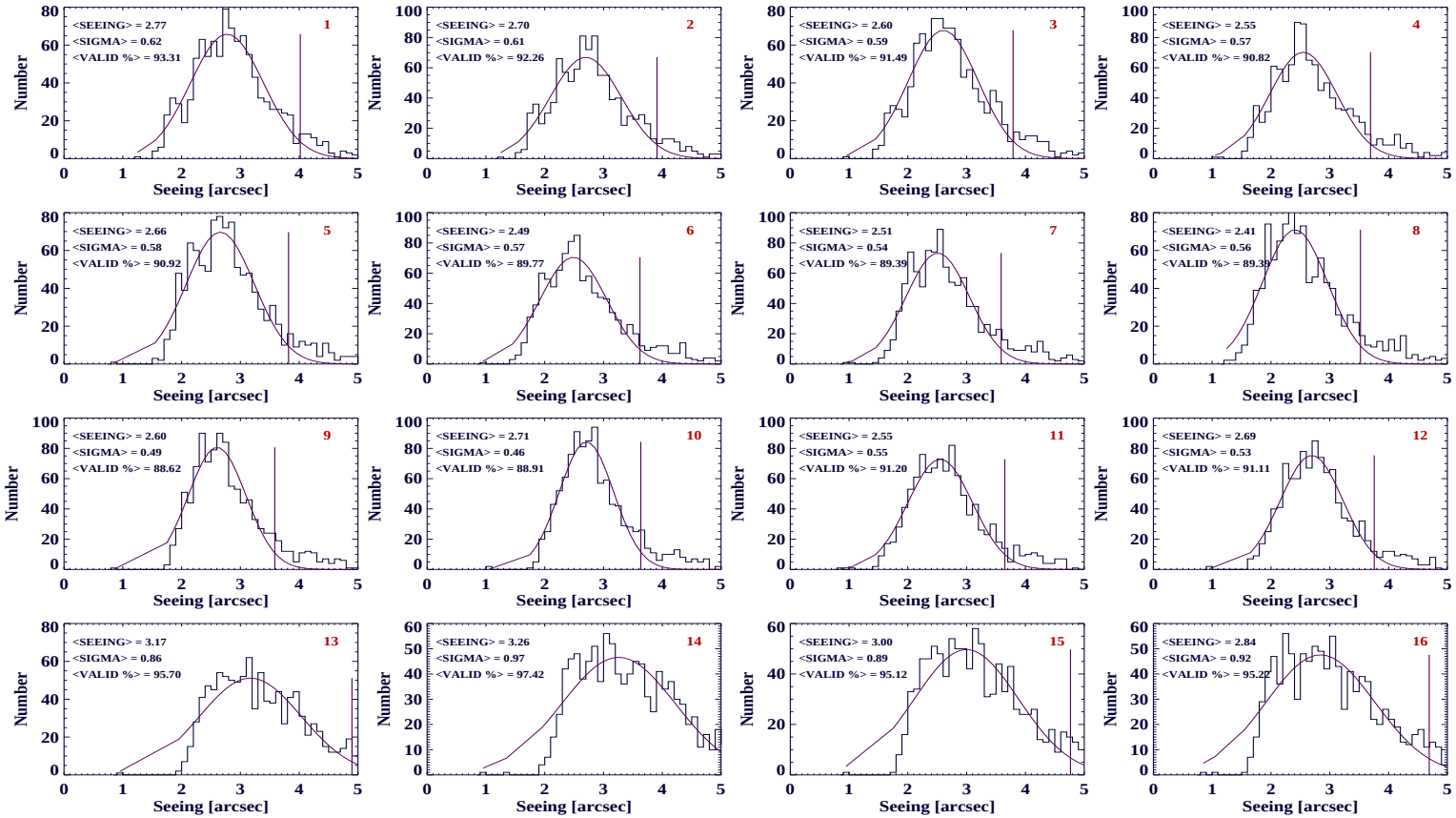

Figure 7.3: Initial seeing distribution of the LAIWOVI field. The curve shows a Gaussian fit to the distribution. The vertical line indicates the 2.0 x σ cut level. The central Gaussian value, σ of the Gaussian and, the remaining percentage of points after a 2.0 x σ cut are also indicated.

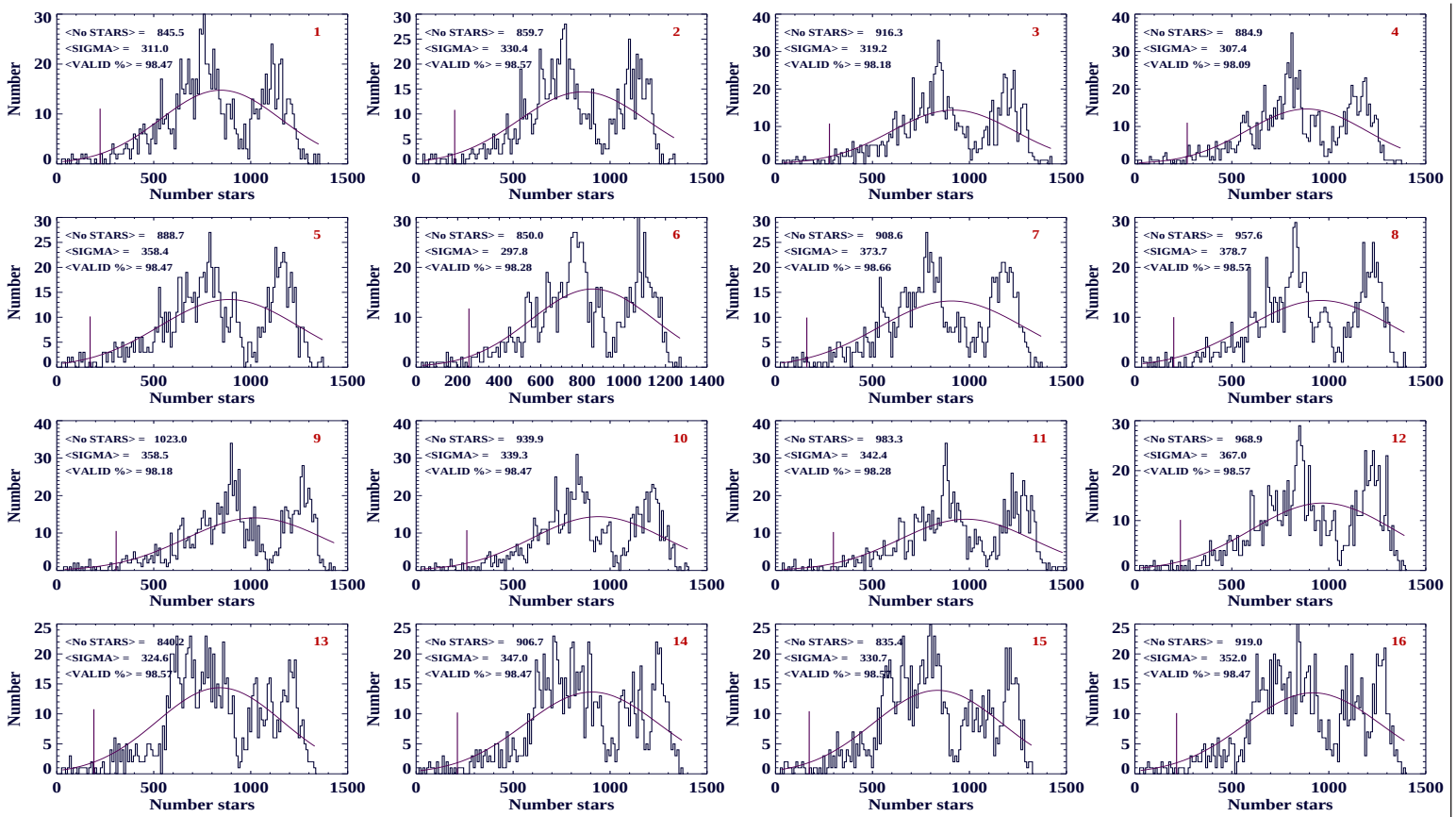

Figure 7.4: Initial number of stars distribution of the LAIWOVI field. The curve shows a Gaussian fit to the distribution. The vertical line indicates the 2.0 x σ cut level. The central Gaussian value, σ of the Gaussian and, the remaining percentage of points after a 2.0 x σ cut are also indicated.

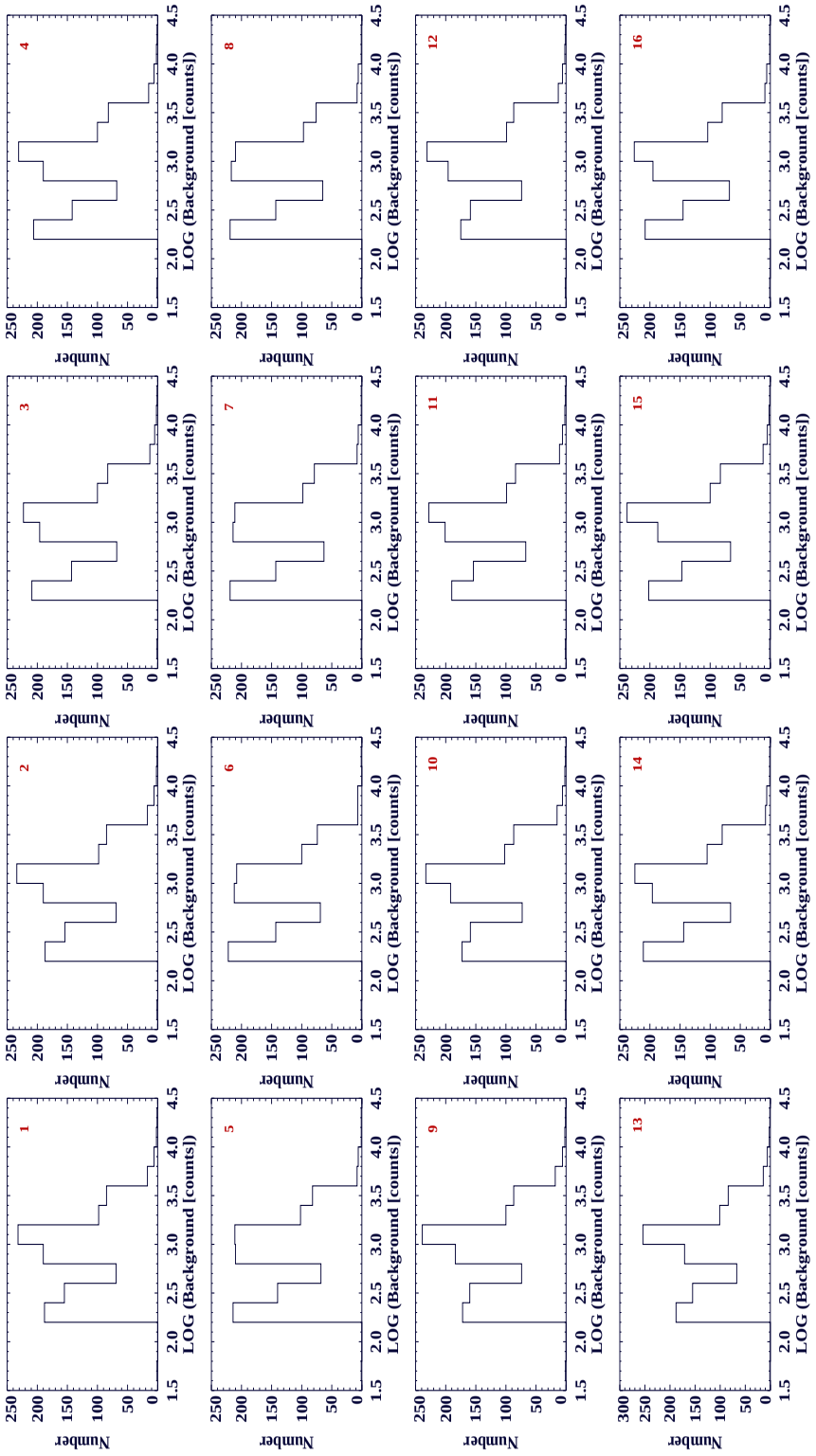

Figure 7.5: Initial background distribution of the LAIWOVI field.

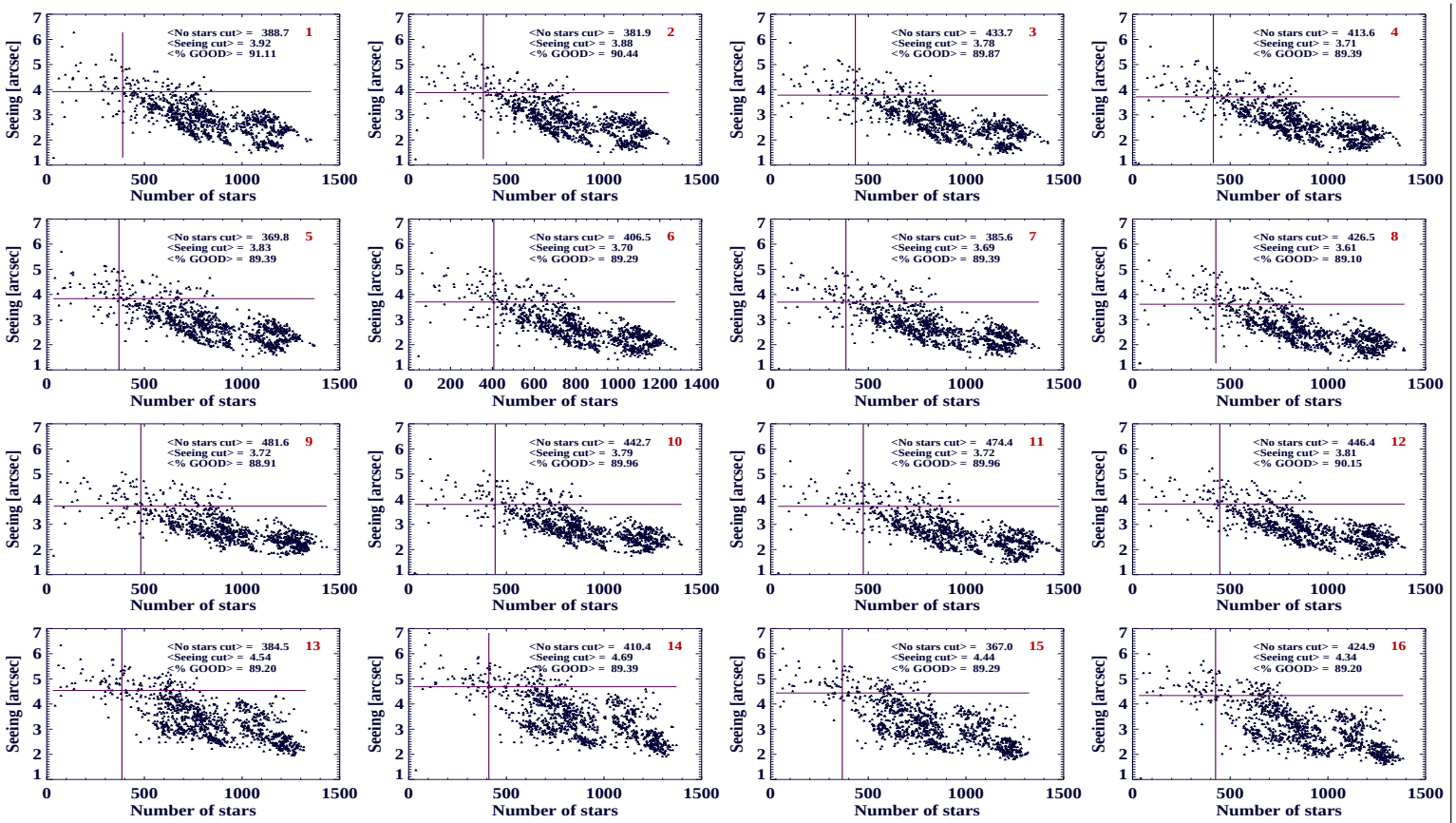

Figure 7.6: Correlation between number of stars and seeing for the field LAIWOVI. The vertical and horizontal lines indicate the initial threshold in number of stars and seeing respectively. The threshold values along with the percentage of good images are also indicated.

7.2.2 Image calibration

The quality of the calibration frames is a key factor to achieve accurate photometry in a transit survey. To avoid introducing spurious variability in this step it is crucial to minimize the systematics in the final light curves. The calibration steps before the photometry are summarized in figure 7.7.

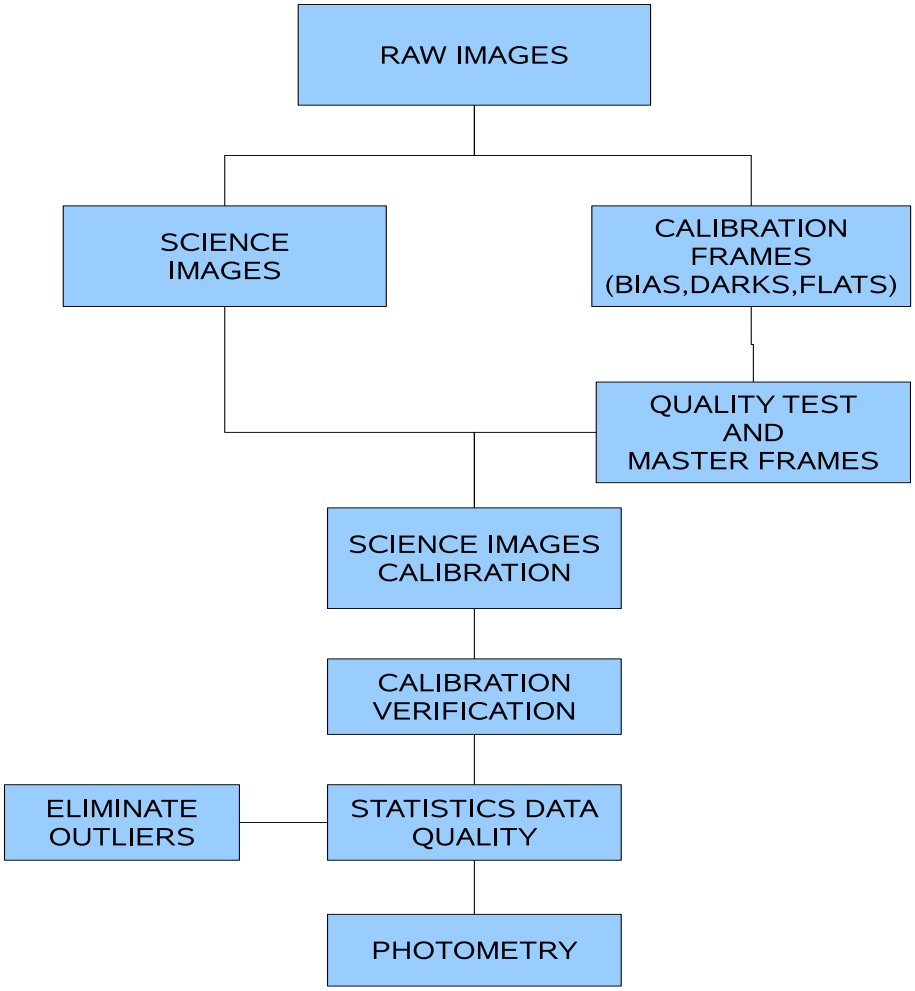

Figure 7.7: Schematic diagram of the LAIWO calibration process.

First, it was necessary to classify the science images and the calibration frames. The light curves were produced in the "R" filter, thus science images and flat fields of only that bandpass were selected. The header information was not accurate, thus we confirmed the coordinates of the target of the desired field for each image. The OBJECT header keyword was compared with

several aliases of the field name and the right ascension and declination were also confirmed.

As a second step, we verified the quality of the BIAS, DARKS, and FLAT FIELDS to minimize their variability and avoid introducing time dependent systematics. A BIAS or DARK frame was deemed good if it had less than 10% variation in each of its quadrants (variability defined as the fractional difference between the median and the mean). A FLAT frame was deemed good if it had less than 10% variation and its median number of counts were enough to represent the pixel to pixel variation accurately (counts between 10,000 and 40,000 in the linear regime of the CCD). The former constrains were imposed in all of the quadrants of a given FLAT. Figure 7.8 shows the fractional variation of the BIAS median as a function of time since 2007 for the first quadrant (1.0 = 100%). As shown in figure 7.8, no significant variations ($<$ 10%) were present. The large variation at $\sim$ 1050 days was due to an anomalous BIAS level on November 11th 2009. In general, BIAS were taken every day, thus, it was possible to calibrate the science images using master BIAS (median combined BIAS) of the same epoch, although its variation was small. Figure 7.9 shows the the fractional variation of the DARK median as a function of time since 2007 for the third quadrant. Variations of the order of $\sim$ 20% were seen, therefore, it was important to calibrate the images with a master DARK (median combined DARK) of the same epoch.

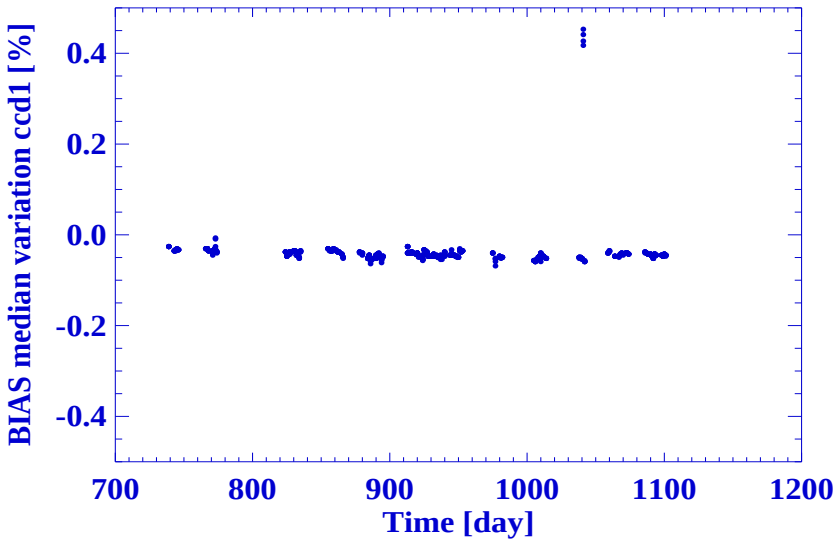

Figure 7.8: Variation of the BIAS median as function of time for the first quadrant.

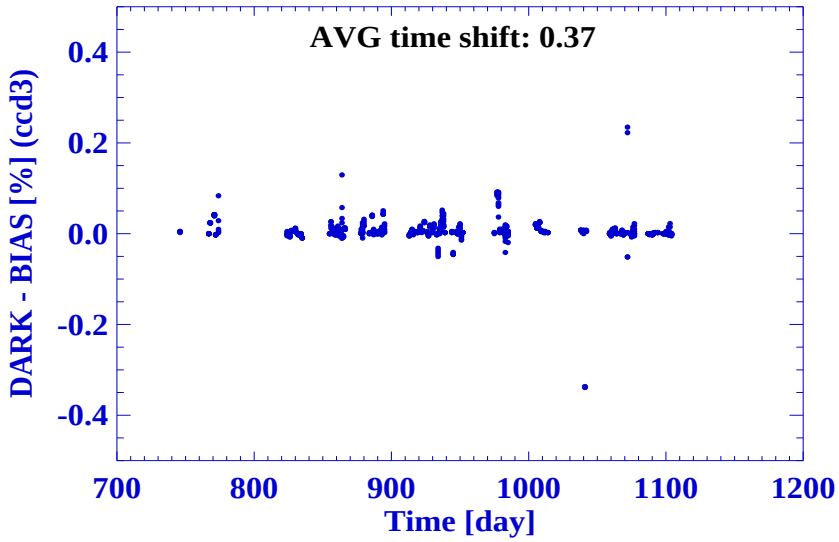

Figure 7.9: Variation of the DARK median as a function of time for the third quadrant. AVG time shift refers to the average time difference between the DARK and the frame used to correct the BIAS level.

As a third step in our analysis, we used IRAF tasks (Tody, 1986) for the image calibration. Once good calibration frames were produced, every science image was BIAS and DARK subtracted and normalized by the FLAT FIELD using master frames of the same day (in some cases this was not possible and the master frames of the closest day were used instead). All the data reduction processes were automated and applied to out large data set. A posterior verification of the calibrated images was run, to ensure that all the images were properly calibrated (BIAS,DARK subtracted and FLAT FIELD normalized).

Finally, preliminary statistics of the images quality were obtained (as it is explained in the previous section). This was done to ensure good images to perform the photometry.

7.3 Light curves

Several steps are required to create good quality light curves. Achieving the desired precision (less than 1.5%) requires the best photometry possible (Udalski et al., 2002a, 2003, 2002b, 2004, 2002c). For example, a Jupiter-like planet in front of a Sun-like star produces an eclipse with a transit depth of only 1%. Moreover, all of the previously discovered transiting extrasolar

planets have transit depths of only a few percents (see exoplanet.eu). To summarize, a good photometric precision is required to detect transiting planets. Our field is not crowded, thus aperture photometry was sufficient. The different steps used to create the light curves are described in the following sections.

7.3.1 Sources detection

We used SeXtractor (Bertin & Arnouts, 1996) for the sources detection. Sources above 3.0 times the RMS and below the saturation level (64,000 counts) were tagged as detections. If the detections were close to the image borders or to each other, they were removed from the list. Detections close to each other ($\leq 6.0 \times pixel$) can blend and cause erroneous photometry measurements. Sources near the image edges ($\leq 12.0 \times pixel$) were most likely not present in all images due to small shifts and also their background estimation is uncertain.

7.3.2 Photometry

We performed aperture photometry. The size of the aperture was chosen according to the background level of the image. The S/N as a function of aperture and background was studied for different magnitudes and the aperture that maximized the S/N was chosen (the apertures range from 3.0 to 4.0 pixel). Before measuring the star flux, the source coordinates were recentered, to ensure that the aperture was centered in every source. Typical coordinate shifts were of the order of $\sim$ 1 pixel. The annulus area was chosen as approximately four times the inner area to guarantee a proper measurement of the background and avoid contamination of neighbour stars. Only stars in the linear regime of the CCD were considered (counts below 35,000 counts), to ensure that the measurements were not affected by instrumental effects. Stars with instrumental magnitudes fainter than 20.0 mag were removed because their measurements were highly uncertain (magnitude errors greater than 1.0 mag).

From the measured stars, the brightest were selected as "flux monitoring stars". They were measured again, but with a big aperture ($\geq 10.0 \times pixel$). They serve to monitor cloud conditions independent of the seeing and they were used to check the relative photometry (explained in section 7.3.7).

7.3.3 Astrometric alignment

Astrometry of every image was computed and corrected. Preliminary astrometry was obtained by comparing the LAIWO field with SDSS-Red images. Then, using the programs SeXtractor,

SCAMP and SWARP (Bertin, 2006; Bertin & Arnouts, 1996) the final astrometric solution was calculated using the 2MASS catalog. Only stars with intermediate S/N (10.0 to 50.0) were used, to avoid using uncertain central positions of bright stars in the astrometric alignment (shifts as big as the field of view $\sim$ 15 arcmin were allowed and used to secure the proper astrometric alignment, although really shifted images were not considered further). The astrometric alignment precision was always of the order of fractions of an arcsec.

7.3.4 Reference image selection

The reference image was chosen as the image with the highest number of detections in the central position. Due to small pointing errors, there were some shifts between images, thus, to maximize the number of points per light curve it was necessary to ensure that the maximum number of stars appeared in all images. Therefore, the image in the central region with the highest number of stars was the adequate choice. Figure 7.10 shows the image distribution on the sky for the field LAIWOVI (first quadrant). Most of the images were contained in the central region. Images within 100 arcsec radius from the central positions were still considered. Images outside this region were not taken into account further because they did not contain most of the stars, therefore several measurements were missing. The typical number of stars of the reference image was about $\sim$ 1200.

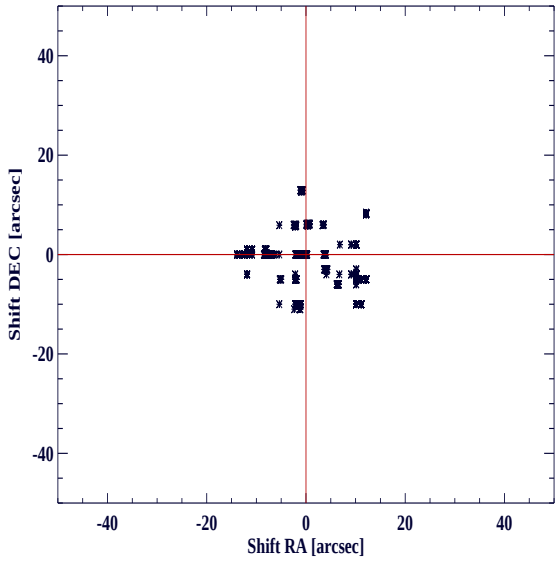

Figure 7.10: Images distribution of the field LAIWOVI for the first quadrant.

Once the reference image was selected, we matched most of its sources with the USNO-A2.0 catalog (Monet, 1998). Figure 7.11 shows the error diagram for the field LAIWOVI first quadrant of the reference image of the matched sources, $\sim$ 20% of the stars per quadrant had photometry precision better than 1.5% (OGLE cut), which is the desired photometric precision to search for transiting planets. Figure 7.12 shows the bright magnitude distribution (R $<$ 16.5) of the field LAIWOVI first quadrant of the reference image. It is seen that most of the stars in the field are faint. Finally, figure 7.13 shows the color magnitude diagram of the field LAIWOVI first quadrant of the reference image.

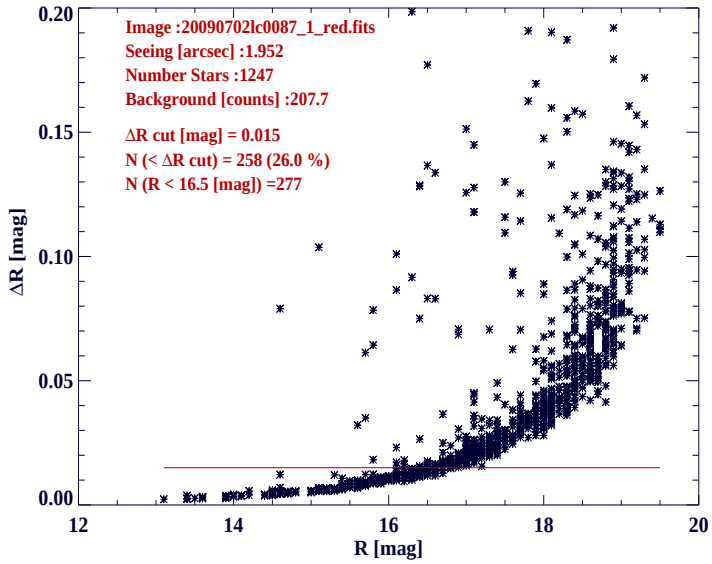

Figure 7.11: Error diagram for the field LAIWOVI first quadrant of the reference image.

7.3.5 Image quality

The image quality was studied again, now taking into account the reference image properties. We performed new cuts in seeing, number of stars and, background. These new cuts were implemented based on the quality of the photometry, which improved without these outliers. We established a common criteria for all quadrants. Table 7.3 shows the boundary values used in the final image quality cut.

Figures 7.14, 7.15 and, 7.16 show the final seeing, number of stars and, background distributions respectively for the first quadrant of the field LAIWOVI.

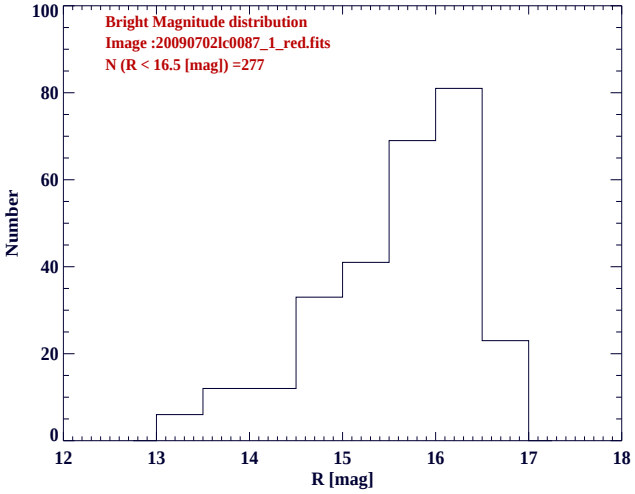

Figure 7.12: Bright magnitudes distribution for the field LAIWOVI of stars of the first quadrant

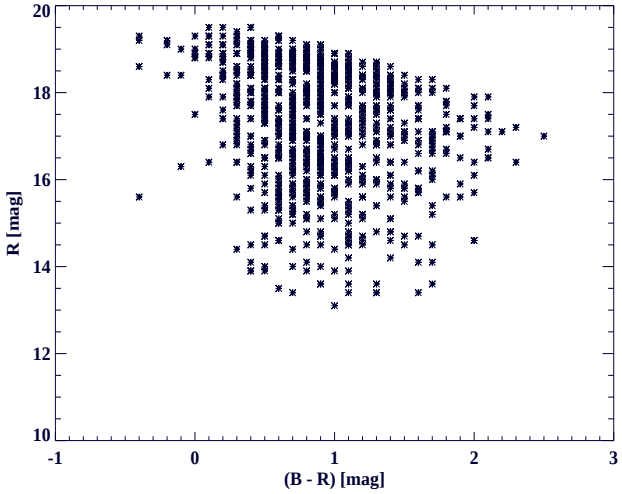

Figure 7.13: Color magnitude diagram for the field LAIWOVI of stars of the first quadrant

Table 7.3: LAIWO image quality cuts

PARAMETER	CUT VALUE
Seeing	3.5 [arcsec]
Number of Stars	(Number stars in reference/4)
Background	12,000 [counts]

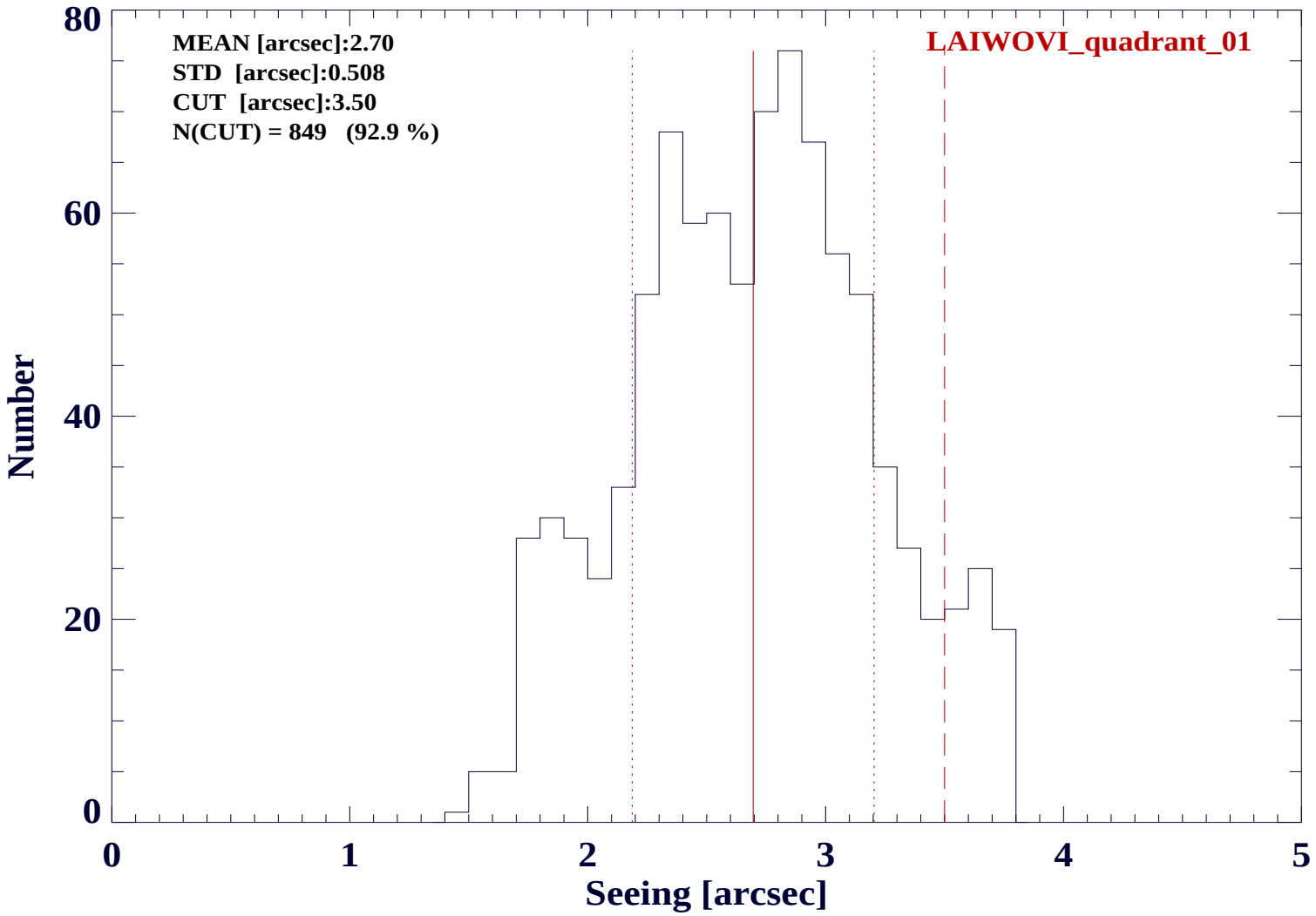

Figure 7.14: Seeing distribution of first quadrant of the field LAIWOVI. The red line indicates the position of the mean, the red dotted lines mark one standard deviation away from the mean, and the dashed line shows the position of the cut value.

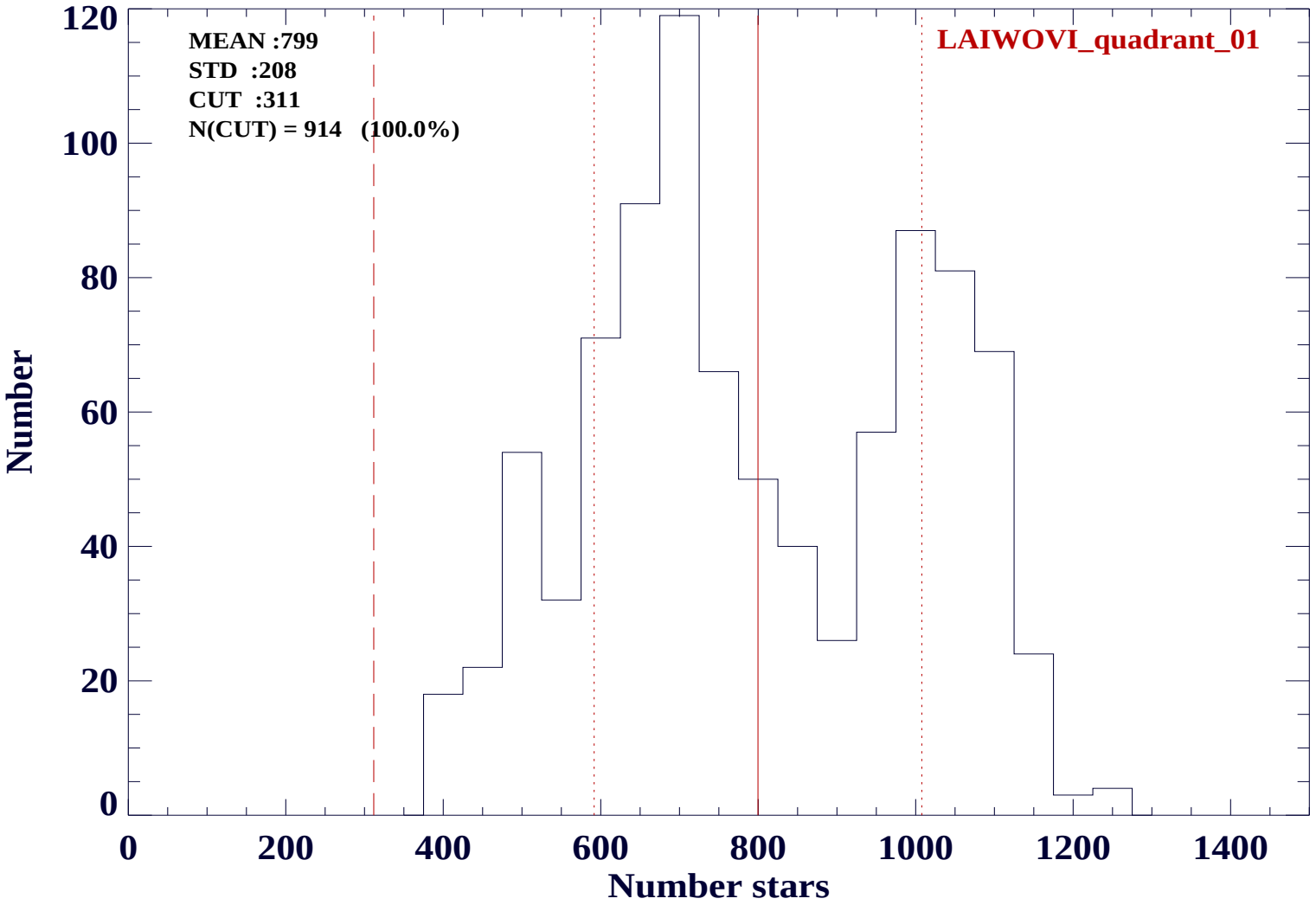

Figure 7.15: Number of stars distribution of first quadrant of the field LAIWOVI. The red line indicates the position of the mean, the red dotted lines mark one standard deviation away from the mean, and the dashed line shows the position of the cut value.

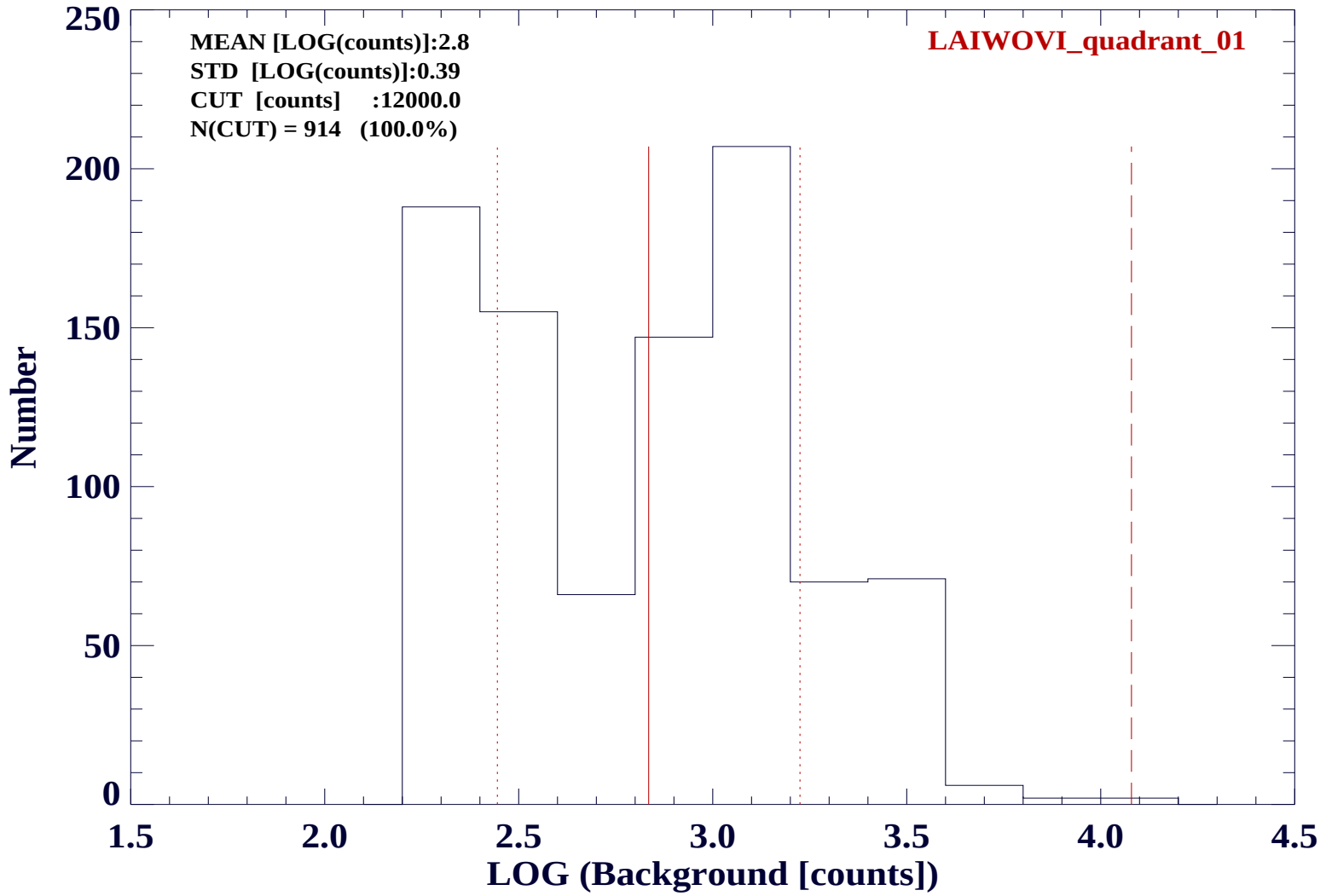

Figure 7.16: Background distribution of first quadrant of the field LAIWOVI. The red line indicates the position of the mean, the red dotted lines mark one standard deviation away from the mean, and the dashed line shows the position of the cut value.

7.3.6 Sources matching

The sources detected in the reference image were defined as the reference stars. All the sources of others images were matched to those stars using the aligned images. The distance between the sources in any image and the reference image sources was calculated and the stars were identified as the closest reference sources. Matched stars with a distance greater than 5.0 [pix] to a reference source or multiple matches for the same star were removed from the list (only the closest source was kept). The typical matching distance was of fraction of pixels.

7.3.7 Relative photometry

The instrumental light curves were calibrated using the method of Scholz & Eislöffel (2004). A set of high quality non-variable reference stars was selected and then the average light curve was computed. This light curve was subtracted from all epochs.

First, an initial sample of potential reference stars was chosen, which had to be present in all epochs with photometric errors below 2.5%.

Second, it was necessary to evaluate the quality of the images. The average instrumental magnitude of every potential reference star was calculated and then this value was subtracted from all time series. The average instrumental magnitude of a star "i" ($i = 1 \ldots N_R$, where N_R is the number of reference stars) is defined as:

$$\bar{m_i} = \frac{1}{N_B} \sum_{j=1}^{N_B} m_i(t_j) \tag{7.1}$$

where N_B is the number of images, and $m_i(t_j)$ is the magnitude of the potential reference star "i" in the image "j". Then, the average subtracted magnitudes of every star are defined as:

$$m_i{}^0(t_j) = m_i(t_j) - \bar{m_i} \tag{7.2}$$

Then, the mean and standard deviations of the average subtracted magnitudes are calculated for each image:

$$\bar{m_j}{}^0 = \frac{1}{N_R} \sum_{i=1}^{N_R} m_i{}^0(t_j) \tag{7.3}$$

$$\sigma_j = \sqrt{\frac{1}{N_R - 1} \sum_{i=1}^{N_R} (m_i{}^0(t_j) - \bar{m_j}{}^0)^2} \tag{7.4}$$

Figure 7.17 shows the last two quantities for the first quadrant of the field LAIWOVI. The mean of the average subtracted magnitudes ($\bar{m_j}^0$) is related to the presence of clouds. This can be seen in figure 7.18. It shows that the changes in atmospheric extinction seen in the flux monitoring star were correlated with this shift, therefore, it permitted to define a cut to eliminate bad quality images with clouds (the measurements of the flux monitoring star were not affected by seeing conditions). The standard deviation of the average subtracted magnitudes (σ_j) is related to the intrinsic image quality. This effect is clearly seen in figure 7.17, where two clumps of points are seen. These evidently indicated the differences in internal quality of images of different epochs. In conclusion, a cut in both, bad quality images with clouds as well as different internal image dispersion helped to improve the photometry quality. Images with some of these two values bigger than 1.5× standard deviation from the mean were deleted (cuts indicated with red dashed lines in figure 7.17).

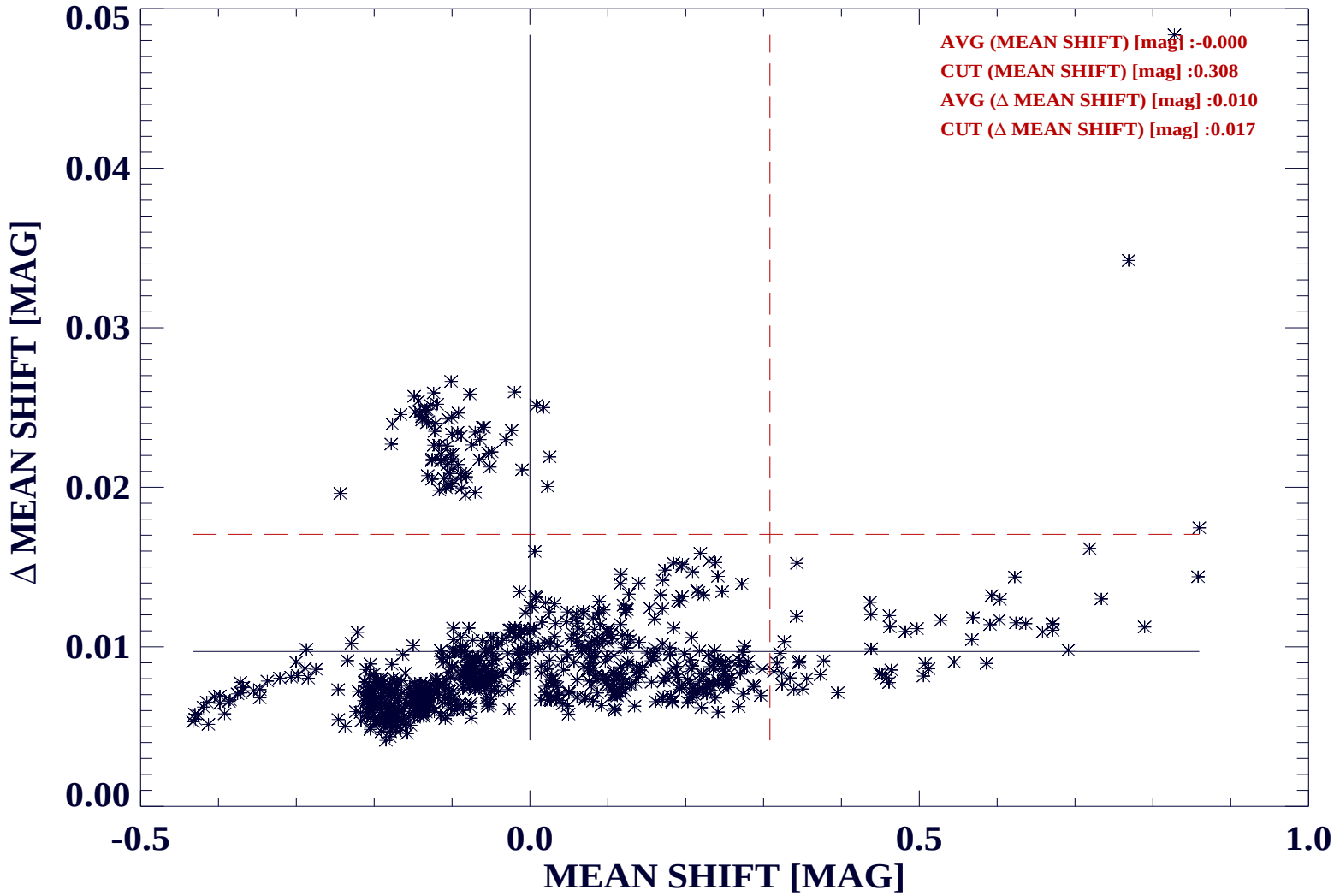

Figure 7.17: Photometry calibration quality for the first quadrant of field LAIWOVI. Black lines show the position of the mean of both quantities. Red dashed lines indicate the position of the cut values. In the upper right corner the important values are shown in red.

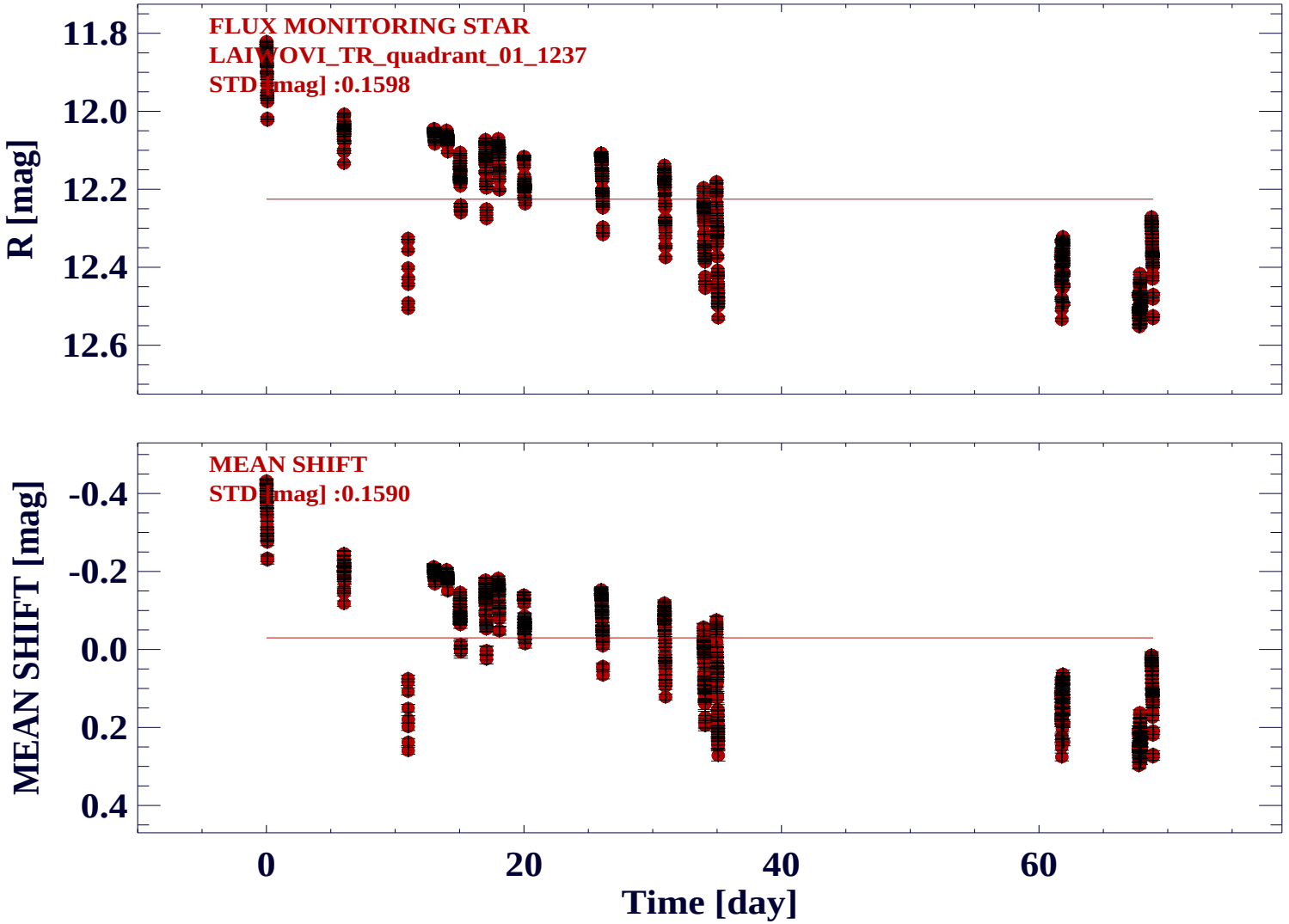

Figure 7.18: Flux monitoring star as a function of mean of the average subtracted magnitudes. It is shown that the mean shift traces the presence of clouds accurately, thus it allows to select the images based on atmospheric extinction.

As a third step, we selected non-variable reference stars. Every star "k" is calibrated using an average light curve of all other stars. This average light curve is defined as:

$$\bar{m_j}' = \sum_{i \neq k} m_i(t_j) \tag{7.5}$$

and the calibrated instrumental magnitudes of the star "k" are:

$$m^0_{k,j} = m_{k,j} - \bar{m_j}' \tag{7.6}$$

The scatter of every potential reference star was calculated as the standard deviation of all $m^0_{k,j}$. Stars with high scatter values were deleted and the process was repeated. The iterations were stopped when the biggest scatter of the reference stars was lower than or equal to 0.8% or the number of stars was lower than or equal to 15, this to avoid reducing the number of stars below 15, that are used to calculate the average light curve. As representative example, from an initial number of 90 good potential reference stars 60 fulfill the described criteria. The mean scatter of those 60 reference stars was 0.6%.

Finally, once the reference stars were chosen, the average reference light curve (also called "the average reference star") was computed as:

$$m^{\bar{REF}}(t_j) = \frac{1}{N_{REF}} \sum_{i=1}^{N_{REF}} m_i{}^{REF}(t_j) \tag{7.7}$$

and it was subtracted from all time series to obtain the extinction corrected relative magnitudes:

$$m^{REL}(t_j) = m(t_j) - m^{\bar{REF}}(t_j) \tag{7.8}$$

Figures 7.19 and 7.20 show two examples where the average reference star was used to retrieve the relative light curve. (both instrumental light curves were used to build the average reference star). In both cases the mean scatter is 0.3%.

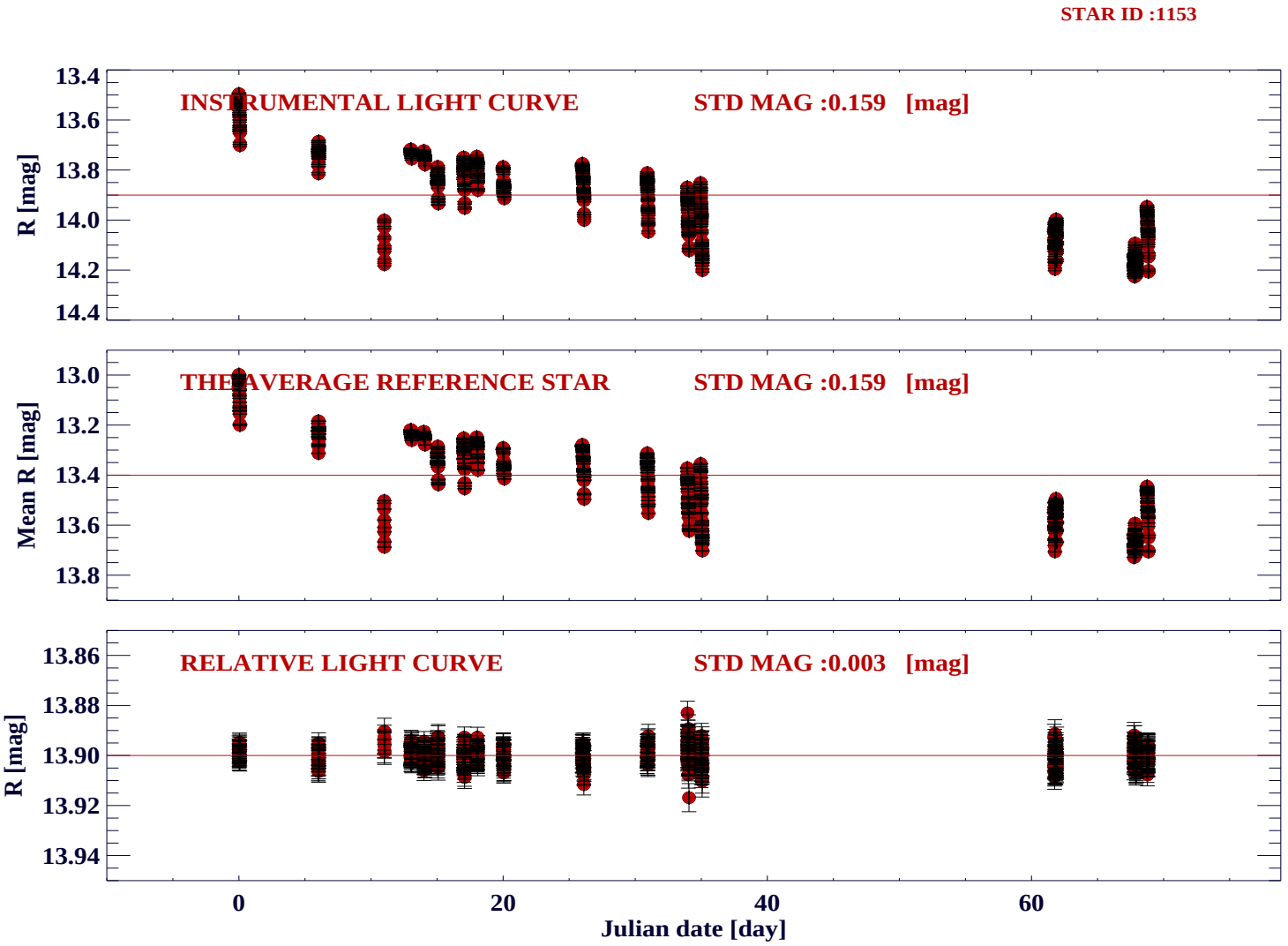

Figure 7.19: Star ID 1153 used to build the "average reference star" for the first quadrant of the field LAIWOVI. The upper panel shows the instrumental light curve, the middle panel the average reference star and the bottom panel the calibrated relative light curve. The mean scatter is 0.3%.

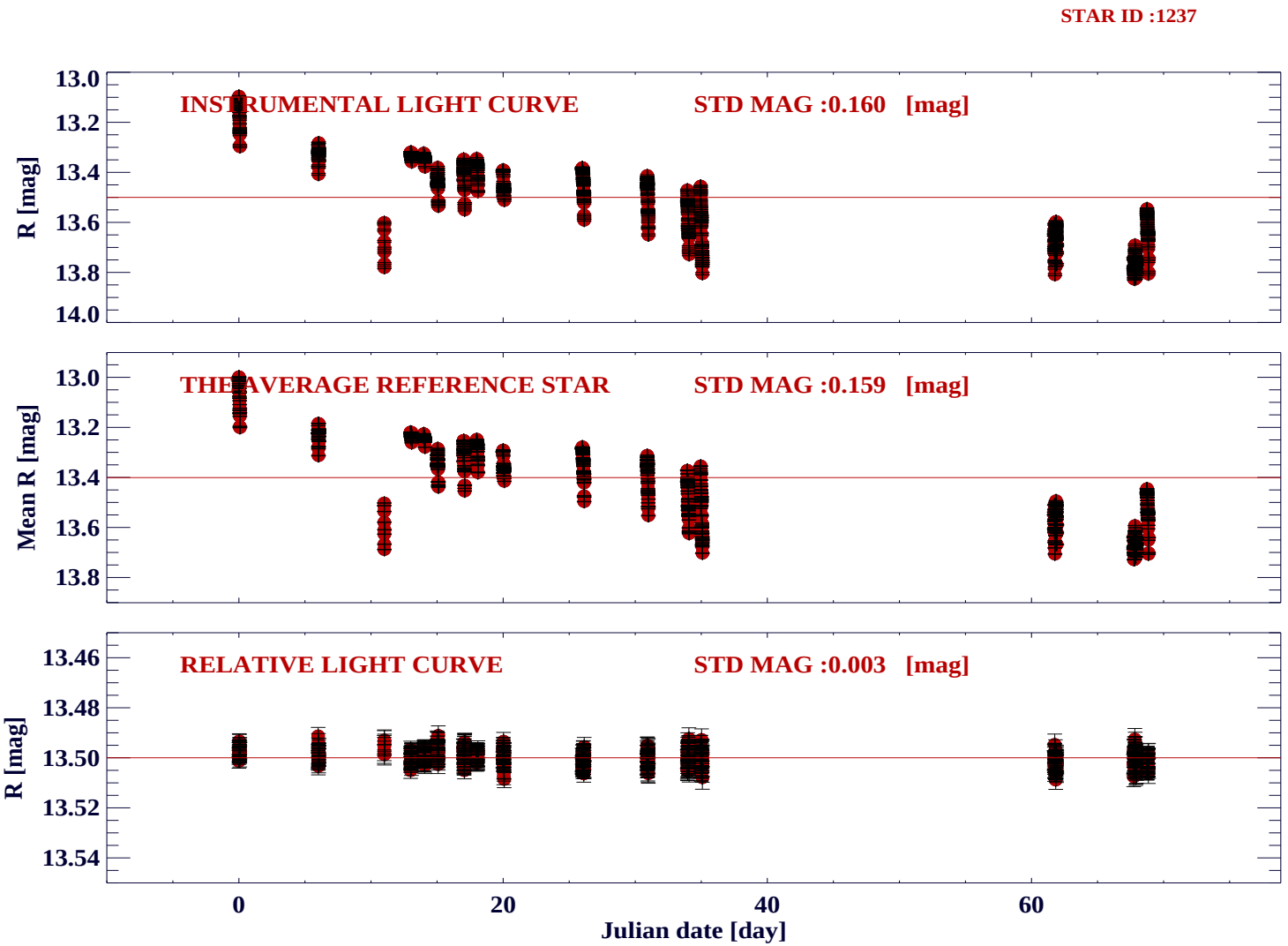

Figure 7.20: Star ID 1237 used to build the "average reference star" for the first quadrant of the field LAIWOVI. The upper panel shows the instrumental light curve, the middle panel the average reference star and the bottom panel the calibrated relative light curve. The mean scatter is 0.3%.

7.3.8 Julian date correction

Due to time errors originated from the telescope, it was necessary to recalculate the Julian dates of the observations again. Shifts of one hour in winter and summer time in Israel introduced uncertainties in the Julian dates of the observations. Therefore, these uncertainties were corrected (as it is explained in the 'Wise Observatory One Meter Telescope Manual', section 5.5, 'Time Stamp') and the julian dates were recalculated.

7.3.9 Summary of cuts

Table 7.4 summarizes the applied cuts and the final number after the cut was done. The relative photometry cut was the strongest constrain (it reduced the number of points in $\sim 15\%$). The final mean number of points for the field LAIWOVI in the 2009 campaign was 690 after all cuts.

Table 7.4: LAIWO cuts summary

QUADRANT	Initial Number [a]	Shift Cut [b]		Image Quality Cut [c]		Photometry Cut [d]	
		Number	Percentage	Number	Percentage	Number	Percentage
quadrant 01	939	914	97.3	849	90.4	717	76.4
quadrant 02	937	913	97.4	853	91.0	732	78.1
quadrant 03	930	906	97.4	882	94.8	718	77.2
quadrant 04	924	900	97.4	890	96.3	766	82.9
quadrant 05	928	904	97.4	880	94.8	721	77.7
quadrant 06	927	904	97.5	894	96.4	729	78.6
quadrant 07	921	899	97.6	896	97.3	728	79.0
quadrant 08	921	899	97.6	899	97.6	714	77.5
quadrant 09	918	898	97.8	887	96.6	732	79.7
quadrant 10	931	909	97.6	882	94.7	766	82.3
quadrant 11	933	910	97.5	902	96.7	749	80.3
quadrant 12	935	912	97.5	885	94.7	804	86.0
quadrant 13	924	894	96.8	651	70.5	577	62.4
quadrant 14	927	897	96.8	605	65.3	490	52.9
quadrant 15	922	895	97.1	706	76.6	538	58.4
quadrant 16	921	894	97.1	756	82.1	583	63.3

[a] It is the initial number of images after the basic image calibration process.

[b] The number after cutting shifted images.

[c] The number after image quality final cuts.

[d] The final number after the relative photometry cuts (intrinsic variability and clouds).

7.4 Systematic Effects

7.4.1 Removal of systematic Effects

We ran the SYSREM algorithm (see section 3.1.2) to remove systematic effects that could have been left in the light curves.

The mean scatter reduction of the light curves as a function of number of corrected effects was studied, i.e. SYSREM was run with different numbers of effects in each case and the improvement of the light curves scatter was studied (see section 3.1.2).

Table 7.5 summarizes the tests. It shows the number of effects that were corrected, the percentage of light curves that were improved and the mean scatter improvement. In all tests, the quality of some light curves was decreased, but in all cases the loss of quality was less than 0.1%. The aim was to reduce the light curves scatter to within 0.5%, i.e. to reduce it to a level comparable to the mean scatter of the reference stars used to build the average light curve. In addition, correlations present in all light curves were also presumably corrected. It was seen that 8 iterations reduced the light curves scatter to within 0.5%, what has been applied to all data.

Table 7.5: SYSREM tests

RUN	No Effects	Good [%]	Scatter Improvement [mag]	Bad[%]	Scatter decreased [mag]
1	2	90.1	-0.0032	9.9	0.0005
2	3	92.5	-0.0036	7.5	0.0006
3	4	94.7	-0.0039	5.3	0.0007
4	5	95.7	-0.0041	4.3	0.0008
5	6	96.8	-0.0044	3.2	0.0006
6	7	97.2	-0.0046	2.8	0.0006
7	8	97.6	-0.0049	2.4	0.0005
8	9	97.8	-0.0052	2.2	0.0002
9	10	98.2	-0.0054	1.8	0.0003
10	11	98.2	-0.0056	1.8	0.0003

Figure 7.21 shows the scatter improvement as a function of magnitude. It can be seen that the improvement was stronger for faint objects than bright objects. This is consistent with the fact that the average light curve was built using bright non-variable stars, thus the bright stars were initially better corrected (similar colors) than the faint stars.

The computation of every coefficient was iterated 20 times to ensure convergence.

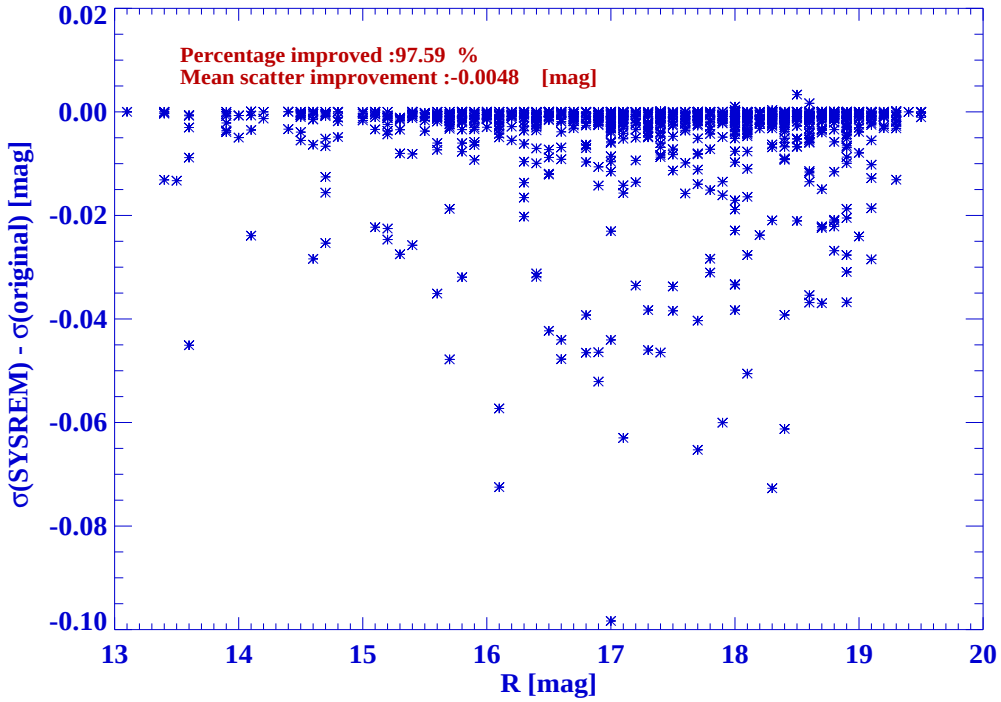

Figure 7.21: Scatter improvement as a function of magnitude for the first quadrant. It can be seen that in most of the stars the scatter was improved. The increase in quality was higher for faint stars than for bright stars (for explanation see text).

7.5 Search for transiting objects

We ran two different approaches to search for transiting objects, one to search for shallower planet-like transits and other to search for eclipsing binaries.

We ran the BLS algorithm (see chapter 4, Kovács et al. 2002) to search for planet-like signatures. Table 7.6 summarizes the parameters used in the BLS run. We used the OGLE criteria (see chapter 5) to define a detection as promissing. Objects with $SDE > 3.0$ and $S/N > 9$ were tagged as detections (for definitions see 4.2). It was also required that the object had at least 3 transits present in the light curve. We decided not to run the TRUFAS algorithm (see chapter 4, Régulo et al. 2007) due to its poor detection efficiency (see section 5.3.2).

We ran the LOMB-SCARGLE periodogram analysis (Lomb, 1976; Scargle, 1982) to search for eclipsing binaries. Table 7.7 summarizes the parameters used in the LOMB-SCARGLE run. Objects with $S/N > 5.0$ (in the periodogram) and with false alarm probability (FAP) less than

Table 7.6: BLS run.

Parameter	Value
Minimum frequency	0.2 $[day^{-1}]$ [a]
Maximum frequency	2.0 $[day^{-1}]$ [b]
Number of frequencies	1501
Number of bins	201
Minimum fractional transit length[c]	0.01
Maximum fractional transit length	0.10

[a] It corresponds to 5.0 [day].

[b] It corresponds to 0.5 [day].

[c] It is defined as transit duration divided by period.

0.01 (99% confidence level) were tagged as promissing.

Table 7.7: LOMB-SCARGLE run.

Parameter	Value
Minimum period	0.1 [day]
Maximum period	5.0 [day]
Period resolution	0.01 [day]

7.5.1 Results

Of the $\sim$ 19,000 stars, a total of 515 candidates were detected by BLS and 4629 by LOMB-SCARGLE periodogram analysis (according to the described criteria).

Visual inspection of the candidates found by both methods gave a total of 18 planet like-transits and 31 eclipsing binaries. It was always necessary to check by eye because noise generated false detections. For example, some of the BLS detections had a period close to 1.0 day, which was the periodicity introduced by the daily gaps of the observations and it was not from astrophysical origin (this was detected also as systematics, but in some cases it was not completely removed).

Tables 7.8 and 7.9 summarize the properties of the discovered transiting planet candidates and eclipsing binaries respectively[1]

[1]We don't quote uncertainties because the detected period could be harmonic/subharmonic.

Table 7.8: Planets candidates found in the LAIWOVI field.

ID	LAIWO ID	COORDINATES (J2000) RA DEC	R [a] [mag]	(B-R) [a]	PERIOD [day]	DEPTH[b] [mag]	DURATION[b] [hr]
1	LAIWOVI_Q1_183.DAT	19 25 27.2 +47 42 15.4	13.60	1.10	2.928	0.0035	6.2
2	LAIWOVI_Q1_542.DAT	19 26 00.8 +47 32 00.0	15.90	0.80	0.929	0.0176	2.1
3	LAIWOVI_Q1_1103.DAT	19 25 09.2 +47 36 10.7	14.60	2.00	1.361	0.0060	3.1
4	LAIWOVI_Q2_347.DAT	19 26 09.0 +47 20 40.0	13.50	0.90	1.706	0.0050	3.9
5	LAIWOVI_Q2_379.DAT	19 26 07.5 +47 21 08.5	14.20	0.90	1.335	0.0041	2.9
6	LAIWOVI_Q3_318.DAT	19 26 49.1 +47 41 12.3	14.20	1.20	2.454	0.0045	4.5
7	LAIWOVI_Q4_120.DAT	19 27 23.4 +47 17 35.8	14.70	1.10	0.886	0.0080	2.0
8	LAIWOVI_Q6_4.DAT	19 25 27.4 +48 11 55.0	14.80	0.90	1.225	0.0053	2.8
9	LAIWOVI_Q6_95.DAT	19 25 28.3 +48 13 17.6	13.90	1.00	2.156	0.0076	4.9
10	LAIWOVI_Q6_525.DAT	19 25 30.4 +48 25 08.8	15.10	0.70	1.241	0.0270	2.7
11	LAIWOVI_Q6_679.DAT	19 24 53.3 +48 17 15.2	15.50	0.60	0.800	0.0173	1.8
12	LAIWOVI_Q8_23.DAT	19 27 22.7 +48 12 55.8	14.10	0.80	1.177	0.0049	2.2
13	LAIWOVI_Q8_93.DAT	19 26 36.6 +48 13 09.1	15.00	0.90	1.177	0.0061	2.7
14	LAIWOVI_Q9_1161.DAT	19 32 25.8 +48 20 04.6	13.20	0.30	0.827	0.0068	1.8
15	LAIWOVI_Q11_146.DAT	19 30 29.4 +48 14 13.9	13.90	1.10	0.947	0.0059	2.1
16	LAIWOVI_Q12_982.DAT	19 30 39.2 +48 30 30.8	15.10	0.90	3.057	0.0160	6.8
17	LAIWOVI_Q16_368.DAT	19 30 24.7 +47 40 20.4	15.40	2.20	1.195	0.0116	2.6
18	LAIWOVI_Q16_1145.DAT	19 30 52.2 +47 36 08.2	14.60	1.10	2.207	0.0108	5.0

[a] From the USNO A2.0 catalog

[b] Parameters assuming a box-shaped transit

Table 7.9: Eclipsing binaries found in the LAIWOVI field.

ID	LAIWO ID	COORDINATES (J2000) RA DEC	R[a] [mag]	(B-R) [a]	Minimum Period [day]
1	LAIWOVI_Q1_211.DAT	19 25 30.9 +47 42 20.9	16.60	0.90	0.272
2	LAIWOVI_Q1_234.DAT	19 26 01.5 +47 42 07.5	16.70	0.40	0.350
3	LAIWOVI_Q1_448.DAT	19 25 53.0 +47 30 41.9	17.80	0.80	0.244
4	LAIWOVI_Q1_580.DAT	19 25 47.1 +47 32 17.6	14.20	1.40	6.716
5	LAIWOVI_Q2_898.DAT	19 25 39.8 +47 26 51.9	17.10	0.70	0.346
6	LAIWOVI_Q3_140.DAT	19 26 47.3 +47 43 41.8	16.80	1.50	0.487
7	LAIWOVI_Q4_365.DAT	19 26 42.8 +47 20 22.2	13.70	0.60	3.666
8	LAIWOVI_Q4_627.DAT	19 27 25.5 +47 28 44.5	18.60	0.00	1.145
9	LAIWOVI_Q6_232.DAT	19 26 04.7 +48 14 39.2	16.40	1.80	1.460
10	LAIWOVI_Q6_359.DAT	19 25 50.1 +48 16 39.9	13.50	1.00	1.884
11	LAIWOVI_Q6_918.DAT	19 25 49.4 +48 18 19.2	17.20	1.50	0.273
12	LAIWOVI_Q6_949.DAT	19 25 28.7 +48 21 24.4	16.30	0.80	0.269
13	LAIWOVI_Q7_1240.DAT	19 27 02.3 +48 34 04.9	17.00	0.70	0.302
14	LAIWOVI_Q8_161.DAT	19 27 05.3 +48 14 03.3	15.00	1.00	2.566
15	LAIWOVI_Q8_528.DAT	19 26 33.3 +48 25 33.5	15.10	2.00	3.902
16	LAIWOVI_Q8_539.DAT	19 26 46.1 +48 25 30.7	18.30	0.10	0.217
17	LAIWOVI_Q9_140.DAT	19 31 51.5 +48 13 27.0	16.10	0.80	0.328
18	LAIWOVI_Q9_166.DAT	19 31 50.6 +48 13 40.1	15.40	1.00	0.945
19	LAIWOVI_Q9_185.DAT	19 32 24.3 +48 13 52.2	16.90	0.70	0.396
20	LAIWOVI_Q9_666.DAT	19 33 04.9 +48 22 52.4	17.00	1.00	0.263
21	LAIWOVI_Q9_826.DAT	19 32 55.1 +48 17 20.7	17.00	0.80	0.331
22	LAIWOVI_Q11_436.DAT	19 31 07.4 +48 16 36.8	15.90	1.00	0.348
23	LAIWOVI_Q11_235.DAT	19 31 02.1 +48 14 56.0	14.20	1.40	2.478
24	LAIWOVI_Q12_124.DAT	19 30 52.7 +48 39 39.0	15.70	1.40	9.454
25	LAIWOVI_Q12_349.DAT	19 30 56.4 +48 36 58.7	15.20	1.50	0.284
26	LAIWOVI_Q13_502.DAT	19 31 50.7 +47 29 01.3	16.30	0.70	0.311
27	LAIWOVI_Q13_798.DAT	19 33 02.3 +47 27 39.3	15.00	0.70	0.821
28	LAIWOVI_Q13_817.DAT	19 31 57.9 +47 27 33.3	3.00	0.00	0.356
29	LAIWOVI_Q14_21.DAT	19 32 29.7 +47 44 19.1	17.30	0.80	0.288
30	LAIWOVI_Q14_181.DAT	19 31 54.8 +47 42 15.8	14.50	0.20	0.844
31	LAIWOVI_Q16_845.DAT	19 30 35.1 +47 33 05.4	17.00	0.80	0.376

[a] From the USNO A2.0 catalog

Figures 7.22 and 7.23 show two examples of positive BLS detections of planet-like transits. In both cases the transit depth is below 1%, therefore, the chance of both transits to be due to a planet is higher. Figures 7.24 and 7.25 show two examples of positive detections of the LOMB-SCARGLE periodogram analysis. In both cases the detected period was a sub-harmonic of the real period (the real period is, in general, bigger than two times the detected period).

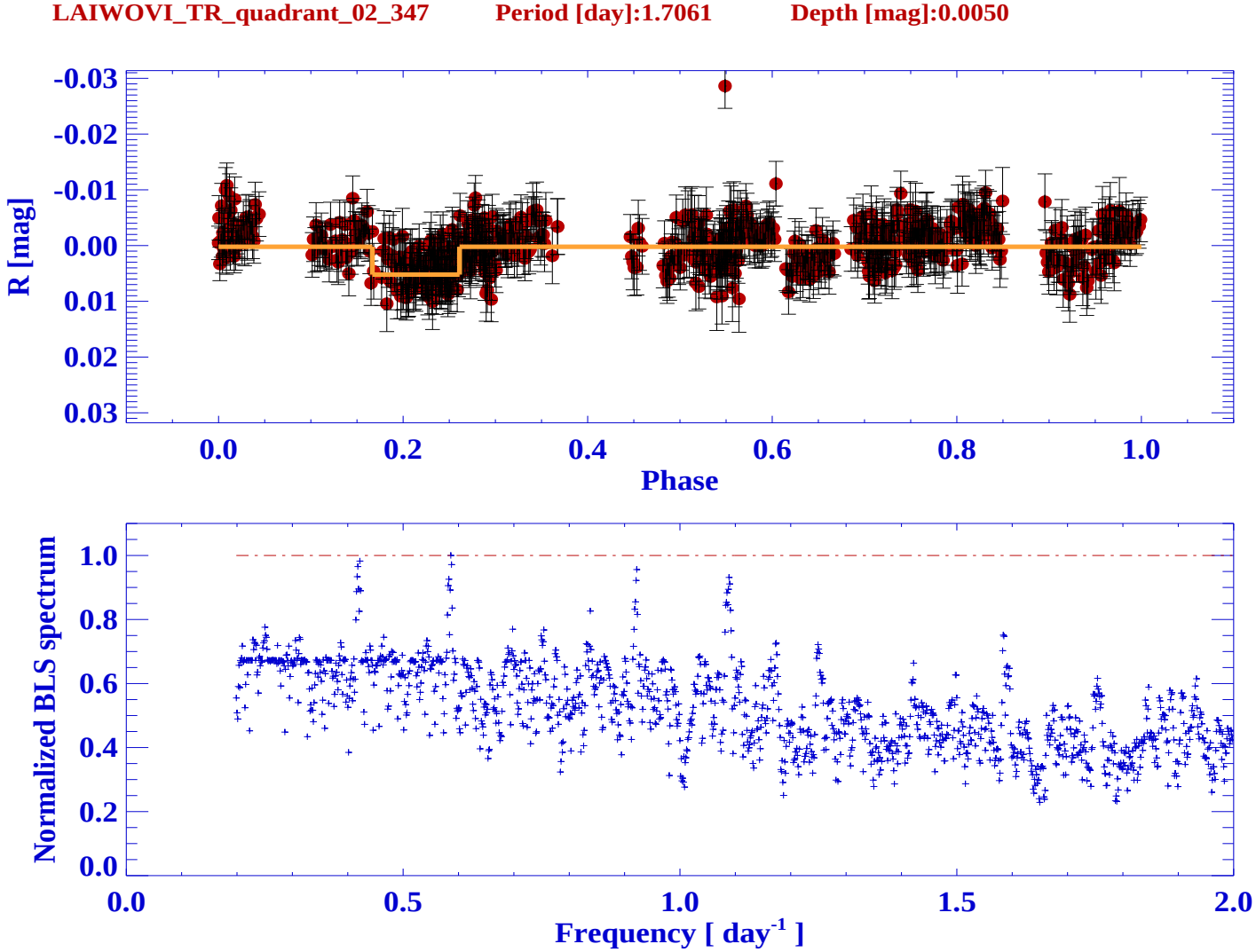

Figure 7.22: Phase-folded light curve and BLS spectrum of the planet-like transit candidate LAIWO-TR-Q2-347. The orange line in the upper panel shows the superimposed BLS fit. The red dashed line in the bottom panel indicates the position of the highest peak of the BLS spectrum. The other prominent peaks represent harmonics/subharmonics of the main detected period.

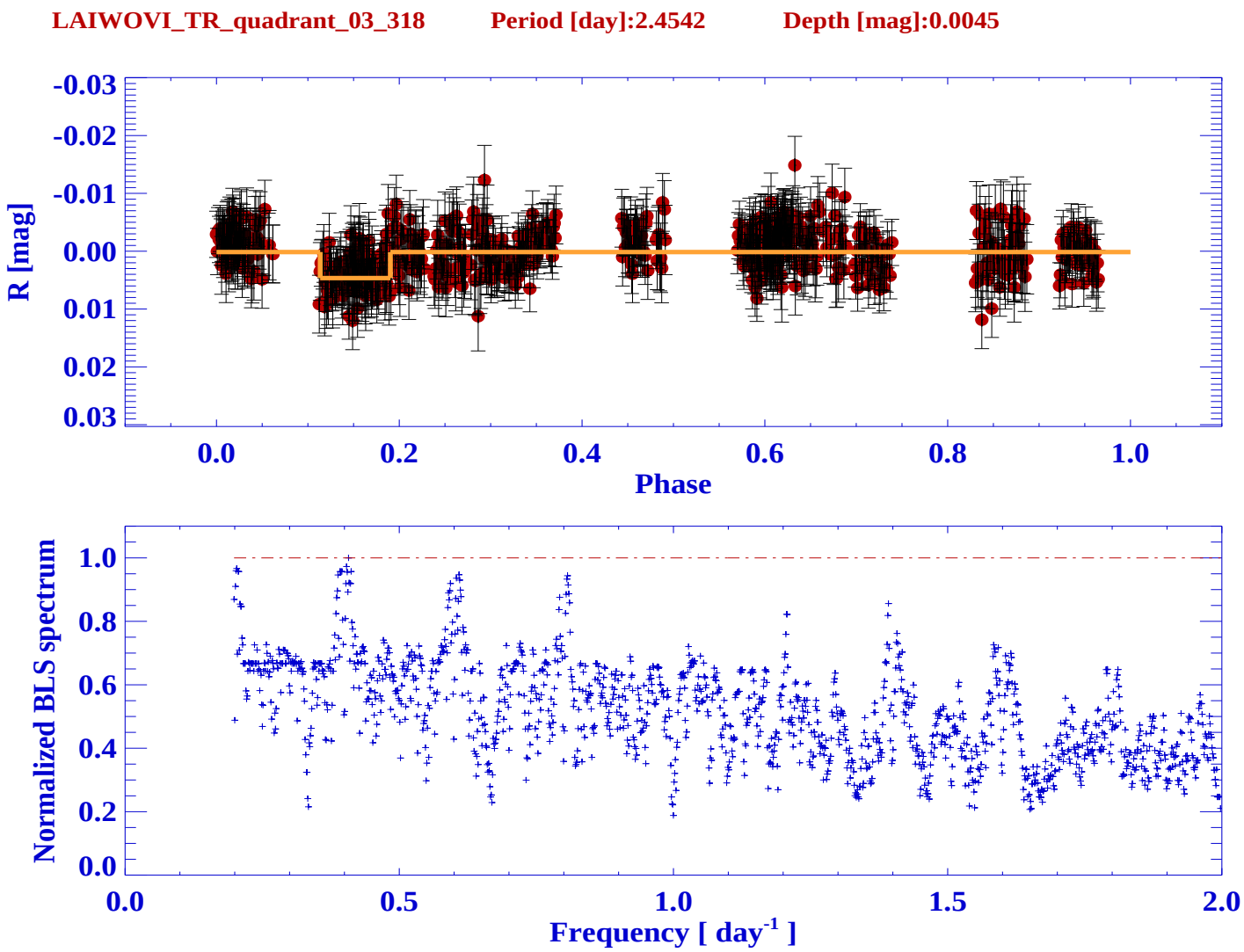

Figure 7.23: Phase-folded light curve and BLS spectrum of the planet-like transit candidate LAIWO-TR-Q3-318. The orange line in the upper panel shows the superimposed BLS fit. The red dashed line in the bottom panel indicates the position of the highest peak of the BLS spectrum. The other prominent peaks represent harmonics/subharmonics of the main detected period.

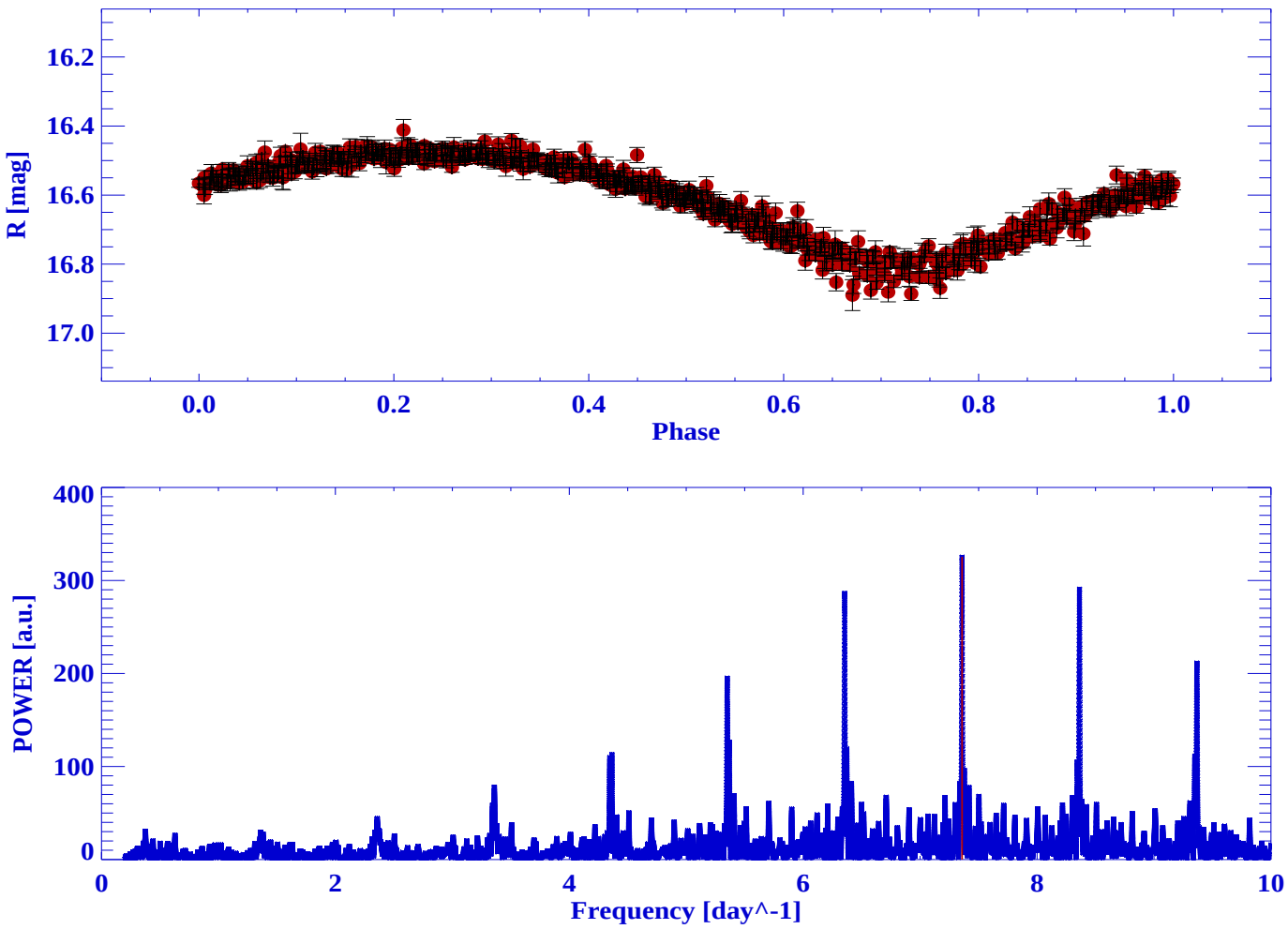

Figure 7.24: Phase folded light curve and Lomb-Scargle periodogram of the eclipsing binary LAIWO-TR-Q1-211. The determined period (in red) is usually a sub-harmonic of the real period, therefore a minimum limit (the value is bigger than twice the detected period). The red line in the bottom panel indicates the position of the found period.

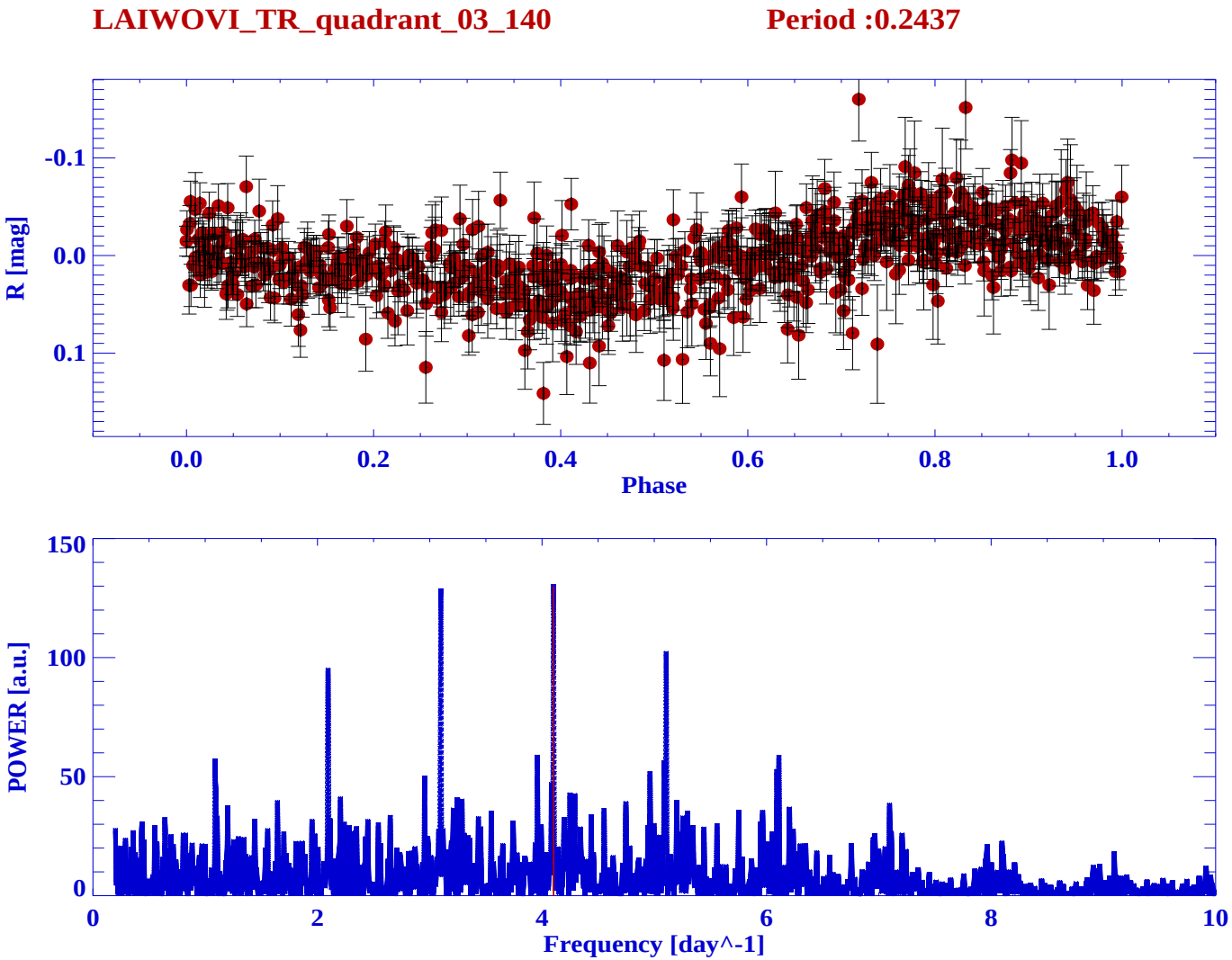

Figure 7.25: Phase folded light curve and Lomb-Scargle periodogram of the eclipsing binary LAIWO-TR-Q3-140. The determined period (in red) is usually a sub-harmonic of the real period, therefore a minimum limit (the value is bigger than twice the detected period) . The red line in the bottom panel indicates the position of the found period.

7.5.2 Analysis

7.5.2.1 Planet Candidates

We have checked if our planet candidates were also found in other surveys. The LAIWOVI field is located in the direction of Cygnus-Lyra region, where the KEPLER survey (Borucki et al., 2009) is monitoring $\sim$ 100,000 main-sequence stars for transiting planets. We have checked our candidate list and found that eight candidates (IDs 4, 5, 7, 12, 13, 16, 17, and 18) had been already studied as candidates by the KEPLER mission [1]. The fact that eight of our candidates were identified previously by the KEPLER mission confirmed that our methodology to build the light curves, remove systematic effects, and the criteria to select the candidates was appropriate and accurate[2]

We have also checked our objects in the *SIMBAD* database [3] and we haven't found any of our candidates identified previously by other surveys. Finally, we have checked the *The Extrasolar Planets Encyclopaedia* [4] and none of our candidates were on the lists of confirmed/suspected transiting extrasolar planets.

We used the color information to determine the spectral type of the host stars if they were main sequence stars (see table 7.11). With the spectral types of the stars we had an approximate stellar radius [5], therefore we could estimate the planet radius of every newly discovered candidate. Table 7.10 summarizes these important properties.

Based on the approximate spectral types of the stars (if they are on the main sequence), the candidates 10, 11, and 14 are too big to be transiting extrasolar planets. Figure 7.26 shows the current planetary radii distribution for the confirmed transiting planets (see exoplanet.eu). It indicates that transiting planets with radii bigger than $1.6 \times R_J$ are rare, therefore the candidates 10, 11, and 14 are most likely not transiting planets. The candidates 2, 3, 6, and 8 are transiting around faint stars ($R > 14.0$), in consequence, their radial velocity confirmation is challenging with the current instrumentation. Candidates 1, 9, and 15 are promising because they orbit solar-like stars and their estimated radii is in the range of the current discovered transiting planets.

[1] For more information about the KEPLER mission visit http://kepler.nasa.gov/
[2] Some of these candidates are still under study and no parameter or data has been released yet.
[3] Website http://simbad.u-strasbg.fr/simbad/
[4] Website http://exoplanet.eu/
[5] See Introduction to Modern Stellar Astrophysics , Appendix G.

Table 7.10: Planet and low mass companion candidate properties

ID	R [mag]	MS Spectral Type[a]	Radius[b] $[R_J]$
1	13.60	G5V	0.6
2	15.90	F8V	1.4
3	14.60	K6V	0.6
6	14.20	G8V	0.6
8	14.80	G0V	0.8
9	13.90	G2V	0.9
10	15.10	F5V	1.9
11	15.50	F2V	1.7
14	13.20	B9V	1.9
15	13.90	G5V	0.7

[a] They were obtained based on the colors of Kraus & Hillenbrand (2007) and are calculated on the assumption of zero extinction.

[b] Planet radii in units of Jupiter radius.

Table 7.11: Colors of main sequence stars

Spectral Type	B	R	(B-R)
F0V	3.00	2.54	0.46
F2V	3.45	2.89	0.56
F5V	3.94	3.23	0.71
F8V	4.52	3.70	0.82
G0V	4.98	4.08	0.90
G2V	5.33	4.35	0.98
G5V	5.78	4.71	1.07
G8V	6.24	5.08	1.16
K0V	6.71	5.43	1.28
K1V	7.01	5.65	1.36
K2V	7.31	5.87	1.44
K3V	7.61	6.08	1.53
K4V	8.06	6.40	1.66
K5V	8.50	6.67	1.83
K7V	9.43	7.27	2.16
M0V	10.2	7.91	2.29
M1V	10.76	8.36	2.40
M2V	11.39	8.90	2.49
M3V	11.91	9.30	2.61
M4V	12.84	10.07	2.77
M5V	13.94	10.82	3.12

[a] See Kraus & Hillenbrand (2007) for more details.

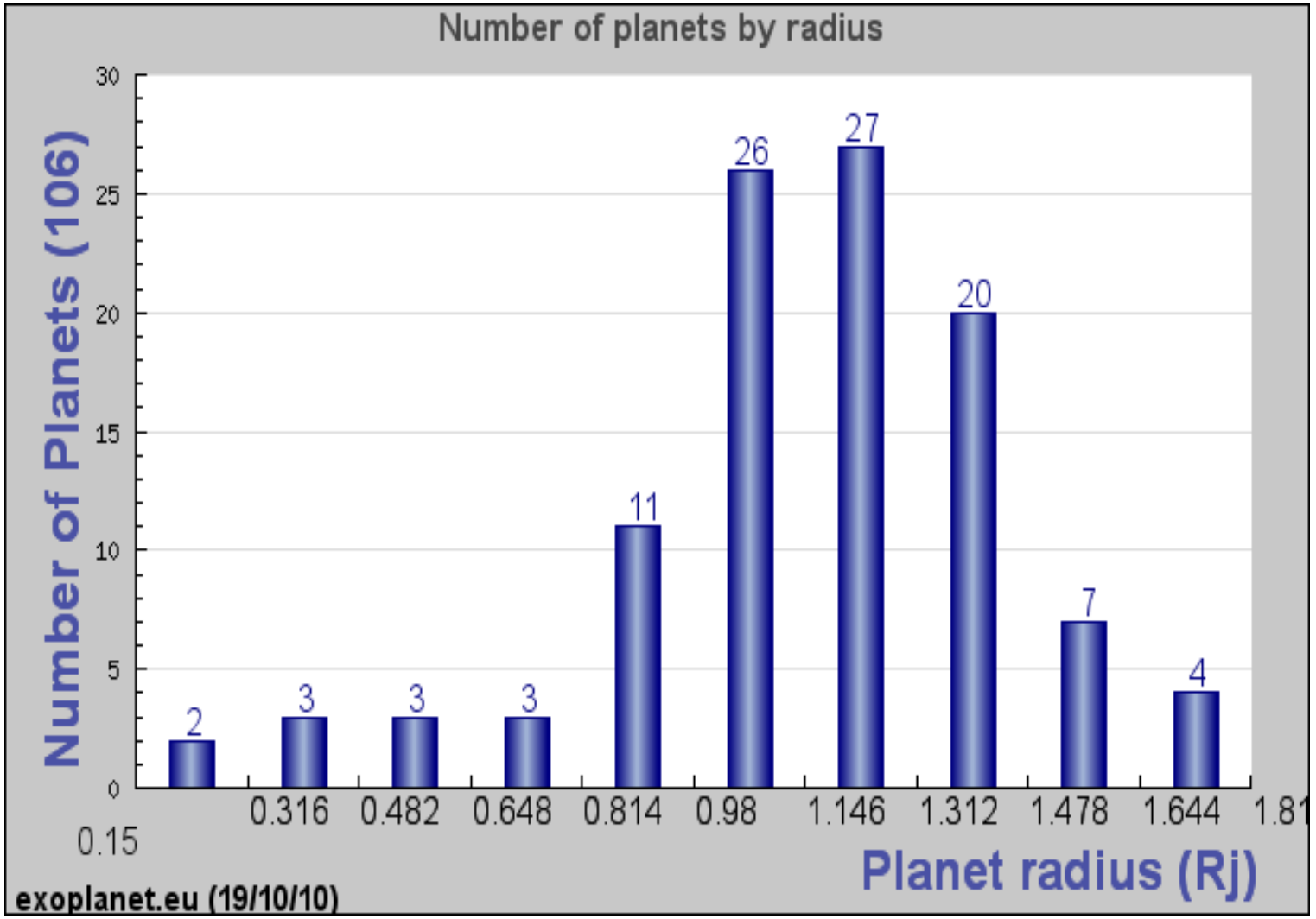

Figure 7.26: Distribution of planetary radii for the confirmed planets (see exoplanet.eu).

7.5.2.2 Eclipsing Binaries

We also checked if our eclipsing binaries were in the KEPLER field. Of the 31 eclipsing binaries, only 3 have been released by the KEPLER survey (IDs 7, 14, and 15).

We have also examined the "General Catalog of Variable Stars" (Samus et al., 2009) [1] and found none of the eclipsing binaries.

In conclusion, we have found 28 eclipsing binaries.

7.6 Conclusions

We have successfully calibrated more than $\sim$ 16,000 images and produced light curves of $\sim$ 19,000 stars.

We have removed the systematic effects which helped to reduce the light curves scatter and to decrease the correlated noise.

We have searched and found 28 new eclipsing binaries.

We have searched and found 3 promising transiting planet candidates (IDs 1, 9, and 15), which fulfill the conditions that are necessary to justify follow-up studies.

The fact the some of our initial candidates are being studied by the KEPLER mission confirms that our methodology and criteria were accurate and correct.

[1] To browse the catalog visit: http://www.sai.msu.su/groups/cluster/gcvs/gcvs/

8

Conclusions

8.1 Conclusions

In this thesis we have studied the necessary steps required to carry out a successful transit search for extrasolar planets.

We have studied the problems that systematics introduce in precise photometry at the millimagnitude level. We have analyzed how to handle and minimize this correlated noise. We have tested the weaknesses and strengths of some transit detection algorithms to search for extrasolar planets using the OGLE data and Pan-Planets simulations. These projects are similar in terms of telescope size and field of view to the LAIWO survey. Finally, we have conducted a transit search for Jupiter-like extrasolar planets using the LAIWO instrument. We have shown how to analyze thousands of images and how to create high quality light curves, which were precise enough to detect promising planet candidates and eclipsing binaries.

In the following sections we summarize the main conclusions and results of the topics covered in this dissertation.

8.1.1 Red Noise

It is important to understand and minimize the systematics in a transit survey. Correlated noise is present at the millimagnitude level, thus complicating the detection of shallower transits that could be due to planets (radial velocity measurements are always necessary to confirm the planetary nature of a companion).

We have quantified the residual red noise in the OGLE data (Udalski et al., 2002a, 2003, 2002b, 2004, 2002c) and Pan-Planets simulations (Koppenhoefer et al., 2009). We have found

that there is always a remaining correlated noise component ($\sim 3mmag$) that must be taken into account. It reduces the significance of a transit, especially for bright stars. For faint stars the photon noise dominates the light curves.

We have minimized the red noise in LAIWO light curves using the SYSREM algorithm (Tamuz et al., 2005). It effectively reduced the scatter of our light curves and helped to detect presumably real transit signals (from astrophysical origin) and not spurious detection due to uncorrected instrumental effects or random noise (for example, without correcting systematics there was a $\sim 1.0day$ periodicity due to the observing duty cycle).

In conclusion, the correlated noise must be reduced and taken into account when evaluating the significance of a transit detection.

8.1.2 Efficiency of the transit detection algorithms

We have tested the two implemented and improved transit detection algorithms (TDA): the Box fitting Least Squares algorithm (BLS, Kovács et al. 2002) and the TRUFAS algorithm (Régulo et al., 2007). We have used the OGLE data set (Udalski et al., 2002a, 2003, 2002b, 2004, 2002c) and Pan-Planets simulations (Koppenhoefer et al., 2009) for these tests.

We have examined the robustness of the BLS algorithm by degrading the OGLE light curves. We have found that the BLS results are robust for this type of ground-based observations. Its average efficiency does not change much (few percents) when the parameters (points in transit, signal to noise, time sampling) experience a big change (50%).

We have evaluated the different survey strategies in the Pan-Planets survey. Using simulations of the first year, we have demonstrated that the TDA provide more reliable results if observational blocks of 1h are used because more transits are present in the light curves and both algorithms are more efficient in that case. Observing 7 fields allows to monitor more stars, and, therefore to discover more transiting planets, while the S/N is still good enough to secure the detections. In conclusion, we have found that the best observational strategy for the Pan-Planets survey that maximizes the detections is the one where 7 fields are monitored with observational blocks of 1h.

The main limitations of the BLS algorithm for these type of ground-based surveys are the transit depth and correlated noise. In both ground-based test data sets (OGLE and Pan-Planets) the BLS effectiveness correlates with these two parameters. The TRUFAS efficiency shows a slightly dependence in the number of points in transit and number of transits present in the light curve (for both ground-based data sets), and, in general its efficiency is low (less than $\sim 50\%$).

Clearly, the BLS algorithm is superior to the TRUFAS algorithm for ground-based observations like OGLE or Pan-Planets.

8.1.3 LAIWO

We have successfully calibrated $\sim$ 16,000 images and produced light curves of $\sim$ 19,000 stars for the "LAIWOVI" field out of which about 26% are bright enough to detect a planet with a transit depth of 1.5%. All the necessary steps to produce these lights curves: *sources detection, photometry, astrometric alignment, reference image selection, image quality test, sources matching, and relative photometry* are automated in a pipeline.

We have minimized the systematic effects using the SYSREM algorithm (Tamuz et al., 2005). We have effectively reduced and eliminated some systematics (like the $\sim$ 1.0 day periodicity of the observations). The detrending of the light curves helped the searching techniques to avoid detecting false random noise due to uncorrected effects.

We have used the BLS algorithm (Kovács et al., 2002) to search for shallower planet-like transits and the LOMB-SCARGLE periodogram analysis (Lomb, 1976; Scargle, 1982) to search for eclipsing binaries. After careful visual inspection of the statistically significant detection of both cases (515 BLS candidates and 4629 LOMB-SCARGLE candidates), we found 18 transits consistent with a planetary companion and 31 eclipsing binaries.

We have checked the existing released candidates of other surveys and found that 8 of our 18 transiting planet candidates are being studied by the Kepler Survey (Borucki et al., 2009), which shows that our reduction processes (calibration of the images, light curves creation, removal of the systematic effects, search for candidates, and criteria to select promising candidates) was accurate and appropriate.

Of the 10 newly discovered transiting planets, 3 are promising to justify follow-up studies. They are orbiting stars that are bright enough and they are smaller than Jupiter. Based on (B-R) colors (assuming zero extinction) the spectral types are G5V, G2V, and G5V if the star in on the main sequence, therefore these stars are suitable to host planets (based on the previous discovered transiting planets).

We have also checked if the eclising binaries were within the KEPLER field and we have found that 3 were already released. Therefore, we have discovered 28 eclipsing binaries.

8.2 Future Work

First, careful analysis of the candidate light curves is necessary to rule out some common false positives, like a massive companion (it is seen as ellipsoidal modulations in the light curve , Sirko & Paczyński 2003). It is also possible to check the transit duration, that should be consistent with a planetary companion (Tingley & Sackett, 2005). If there are no indications of a massive companion, the next step is high resolution spectroscopy.

High resolution spectroscopy and radial velocity follow-up are crucial to confirm or reject the planetary nature of a transiting companion. Radial velocity measurements allow the determination of the mass of the transiting object, which is the only way to discriminate planets from low-mass stars or brown dwarfs. These objects have sizes similar to Jupiter, therefore they produce the same eclipse light curve. High resolution spectroscopic is the only way to secure the planetary nature, characterize the properties of the host star and confirm the mass the transiting object (Bouchy et al., 2005a; Bouchy & Queloz, 2007).

Various indicators in the high resolution spectra would reveal the presence of a massive companion (Bouchy & Queloz, 2007): multi-component spectra (the Cross Correlation Function of the spectrum exhibits more than one component), massive tidally locked companion (stellar rotation period is synchronized, i.e. $P_{rot} = P_{transit}$), spectral type incompatible with planetary companion (stellar radius indicates the size of the transiting object outside planetary regime).

A large radial velocity variation indicates unambiguously an object with mass in the stellar regime. For a circular orbit, the semi-amplitude of the radial velocity variation (in kms^{-1}) is directly related to the mass (in solar units) and period (in days) of the companion through:

$$K = 214 \times \frac{m}{(m+M)^{2/3}} \times P^{-1/3}$$

A $0.1 \times M_{\odot}$ star with a $10.0 \times M_J$ companion orbiting with a 10 days orbit gives $K = 1.0 kms^{-1}$. In general, if the semi-amplitude is larger than few kms^{-1}, the companion is clearly not in the planetary regime.

Bibliography

Afonso, C. & Henning, T. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 366, Transiting Extrapolar Planets Workshop, ed. C. Afonso, D. Weldrake, & T. Henning, 326–+ 59

Alcock, C., Allsman, R., Alves, D. R., Axelrod, T., Becker, A., Bennett, D., Clement, C., Cook, K. H., Drake, A., Freeman, K., Geha, M., Griest, K., Kovács, G., Kurtz, D. W., Lehner, M., Marshall, S., Minniti, D., Nelson, C., Peterson, B., Popowski, P., Pratt, M., Quinn, P., Rodgers, A., Rowe, J., Stubbs, C., Sutherland, W., Tomaney, A., Vandehei, T., & Welch, D. L. 2000, Astrophys. J, 542, 257 29

Baglin, A. 2003, Advances in Space Research, 31, 345 25, 32

Baglin, A., Auvergne, M., Barge, P., Buey, J., Catala, C., Michel, E., Weiss, W., & COROT Team. 2002, in ESA Special Publication, Vol. 485, Stellar Structure and Habitable Planet Finding, ed. B. Battrick, F. Favata, I. W. Roxburgh, & D. Galadi, 17–24 6

Bakos, G., Noyes, R. W., Kovács, G., Stanek, K. Z., Sasselov, D. D., & Domsa, I. 2004, Publ.. Astron. Soc. Pac, 116, 266 6

Baumeister, H., Afonso, C., Marien, K., & Klein, R. 2006, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 6269, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series 87

Bertin, E. 2006, in Astronomical Society of the Pacific Conference Series, Vol. 351, Astronomical Data Analysis Software and Systems XV, ed. C. Gabriel, C. Arviset, D. Ponz, & S. Enrique, 112–+ 99

Bertin, E. & Arnouts, S. 1996, Astron. Astrophys. Suppl. Ser, 117, 393 98, 99

BIBLIOGRAPHY

Bond, I. A., Udalski, A., Jaroszyński, M., Rattenbury, N. J., Paczyński, B., Soszyński, I., Wyrzykowski, L., Szymański, M. K., Kubiak, M., Szewczyk, O., Żebruń, K., Pietrzyński, G., Abe, F., Bennett, D. P., Eguchi, S., Furuta, Y., Hearnshaw, J. B., Kamiya, K., Kilmartin, P. M., Kurata, Y., Masuda, K., Matsubara, Y., Muraki, Y., Noda, S., Okajima, K., Sako, T., Sekiguchi, T., Sullivan, D. J., Sumi, T., Tristram, P. J., Yanagisawa, T., & Yock, P. C. M. 2004, Astrophys. J., Lett, 606, L155 2

Borucki, W., Koch, D., Batalha, N., Caldwell, D., Christensen-Dalsgaard, J., Cochran, W. D., Dunham, E., Gautier, T. N., Geary, J., Gilliland, R., Jenkins, J., Kjeldsen, H., Lissauer, J. J., & Rowe, J. 2009, in IAU Symposium, Vol. 253, IAU Symposium, 289–299 89, 124, 131

Borucki, W. J., Koch, D., Basri, G., Batalha, N., Brown, T., Caldwell, D., Caldwell, J., Christensen-Dalsgaard, J., Cochran, W. D., DeVore, E., Dunham, E. W., Dupree, A. K., Gautier, T. N., Geary, J. C., Gilliland, R., Gould, A., Howell, S. B., Jenkins, J. M., Kondo, Y., Latham, D. W., Marcy, G. W., Meibom, S., Kjeldsen, H., Lissauer, J. J., Monet, D. G., Morrison, D., Sasselov, D., Tarter, J., Boss, A., Brownlee, D., Owen, T., Buzasi, D., Charbonneau, D., Doyle, L., Fortney, J., Ford, E. B., Holman, M. J., Seager, S., Steffen, J. H., Welsh, W. F., Rowe, J., Anderson, H., Buchhave, L., Ciardi, D., Walkowicz, L., Sherry, W., Horch, E., Isaacson, H., Everett, M. E., Fischer, D., Torres, G., Johnson, J. A., Endl, M., MacQueen, P., Bryson, S. T., Dotson, J., Haas, M., Kolodziejczak, J., Van Cleve, J., Chandrasekaran, H., Twicken, J. D., Quintana, E. V., Clarke, B. D., Allen, C., Li, J., Wu, H., Tenenbaum, P., Verner, E., Bruhweiler, F., Barnes, J., & Prsa, A. 2010, Science, 327, 977 6

Bouchy, F., Pont, F., Melo, C., Santos, N. C., Mayor, M., Queloz, D., & Udry, S. 2005a, Astron. Astrophys, 431, 1105 132

Bouchy, F., Pont, F., Santos, N. C., Melo, C., Mayor, M., Queloz, D., & Udry, S. 2004, Astron. Astrophys, 421, L13 29, 30, 34, 35, 39

Bouchy, F. & Queloz, D. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 366, Transiting Extrapolar Planets Workshop, ed. C. Afonso, D. Weldrake, & T. Henning, 193–+ 132

Bouchy, F., Udry, S., Mayor, M., Moutou, C., Pont, F., Iribarne, N., da Silva, R., Ilovaisky, S., Queloz, D., Santos, N. C., Ségransan, D., & Zucker, S. 2005b, Astron. Astrophys, 444, L15 11, 12

Brown, T. M. 2003, Astrophys. J., Lett, 593, L125 8

Brown, T. M. & Charbonneau, D. 2000, in Astronomical Society of the Pacific Conference Series, Vol. 219, Disks, Planetesimals, and Planets, ed. G. Garzón, C. Eiroa, D. de Winter, & T. J. Mahoney, 584–+ 6

Charbonneau, D., Allen, L. E., Megeath, S. T., Torres, G., Alonso, R., Brown, T. M., Gilliland, R. L., Latham, D. W., Mandushev, G., O'Donovan, F. T., & Sozzetti, A. 2005, Astrophys. J, 626, 523 11

Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. 2000, Astrophys. J., Lett, 529, L45 6, 7, 11

Charbonneau, D., Brown, T. M., Noyes, R. W., & Gilliland, R. L. 2002, Astrophys. J, 568, 377 11

Charbonneau, D., Knutson, H. A., Barman, T., Allen, L. E., Mayor, M., Megeath, S. T., Queloz, D., & Udry, S. 2008, Astrophys. J, 686, 1341 12, 13, 14

Chyba, C. F. & Hand, K. P. 2005, Ann. Rev. Astron. Astrophys, 43, 31 1

Defaÿ, C., Deleuil, M., & Barge, P. 2001, Astron. Astrophys, 365, 330 42, 64, 70

Deming, D., Seager, S., Richardson, L. J., & Harrington, J. 2005, Nature, 434, 740 11

Drake, A. J. 2003, Astrophys. J, 589, 1020 42

Gorbikov, E., Brosch, N., & Afonso, C. 2010, Astrophys. Space. Sci, 326, 203 88

Guillot, T. 2005, Annual Review of Earth and Planetary Sciences, 33, 493 7

Holman, M. J., Fabrycky, D. C., Ragozzine, D., Ford, E. B., Steffen, J. H., Welsh, W. F., Lissauer, J. J., Latham, D. W., Marcy, G. W., Walkowicz, L. M., Batalha, N. M., Jenkins, J. M., Rowe, J. F., Cochran, W. D., Fressin, F., Torres, G., Buchhave, L. A., Sasselov, D. D., Borucki, W. J., Koch, D. G., Basri, G., Brown, T. M., Caldwell, D. A., Charbonneau, D., Dunham, E. W., Gautier, T. N., Geary, J. C., Gilliland, R. L., Haas, M. R., Howell, S. B., Ciardi, D. R., Endl, M., Fischer, D., Fürész, G., Hartman, J. D., Isaacson, H., Johnson, J. A., MacQueen, P. J., Moorhead, A. V., Morehead, R. C., & Orosz, J. A. 2010, Science, 330, 51 15, 16

Holman, M. J. & Murray, N. W. 2005, Science, 307, 1288 13, 15

Kalas, P., Graham, J. R., Chiang, E., Fitzgerald, M. P., Clampin, M., Kite, E. S., Stapelfeldt, K., Marois, C., & Krist, J. 2008, Science, 322, 1345 2

Konacki, M., Torres, G., Jha, S., & Sasselov, D. D. 2003, Nature, 421, 507 6, 39, 41

Konacki, M., Torres, G., Sasselov, D. D., & Jha, S. 2005, Astrophys. J, 624, 372 39

Koppenhoefer, J., Afonso, C., Saglia, R. P., & Henning, T. 2009, Astron. Astrophys, 494, 707 21, 59, 60, 80, 129, 130

Kovács, G., Bakos, G., & Noyes, R. W. 2005, Mon. Not. R. Astron. Soc, 356, 557 19

Kovács, G., Zucker, S., & Mazeh, T. 2002, Astron. Astrophys, 391, 369 25, 26, 27, 31, 52, 58, 63, 115, 130, 131

Kraus, A. L. & Hillenbrand, L. A. 2007, Astrophysical Journal, 134, 2340 125, 126

Kruszewski, A. & Semeniuk, I. 2003, Acta Astronomica, 53, 241 18, 41

Lomb, N. R. 1976, Astrophys. Space. Sci, 39, 447 115, 131

Mao, S. & Paczynski, B. 1991, Astrophys. J., Lett, 374, L37 2

Mayor, M. & Queloz, D. 1995, Nature, 378, 355 1, 6

Mazeh, T., Tamuz, O., & Zucker, S. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 366, Transiting Extrapolar Planets Workshop, ed. C. Afonso, D. Weldrake, & T. Henning, 119–+ 19

McCullough, P. R., Stys, J. E., Valenti, J. A., Fleming, S. W., Janes, K. A., & Heasley, J. N. 2005, Publ.. Astron. Soc. Pac, 117, 783 6

McLaughlin, D. B. 1924, Astrophys. J, 60, 22 12

Monet, D. G. 1998, in Bulletin of the American Astronomical Society, Vol. 30, Bulletin of the American Astronomical Society, 1427–+ 100

Monet, D. G., Dahn, C. C., Vrba, F. J., Harris, H. C., Pier, J. R., Luginbuhl, C. B., & Ables, H. D. 1992, Astrophysical Journal, 103, 638 18

Moutou, C., Hébrard, G., Bouchy, F., Eggenberger, A., Boisse, I., Bonfils, X., Gravallon, D., Ehrenreich, D., Forveille, T., Delfosse, X., Desort, M., Lagrange, A., Lovis, C., Mayor, M., Pepe, F., Perrier, C., Pont, F., Queloz, D., Santos, N. C., Ségransan, D., Udry, S., & Vidal-Madjar, A. 2009, Astron. Astrophys, 498, L5 6

Moutou, C., Pont, F., Barge, P., Aigrain, S., Auvergne, M., Blouin, D., Cautain, R., Erikson, A. R., Guis, V., Guterman, P., Irwin, M., Lanza, A. F., Queloz, D., Rauer, H., Voss, H., & Zucker, S. 2005, Astron. Astrophys, 437, 355 25, 33, 34

Moutou, C., Pont, F., & Halbwachs, J. 2006, Formation planétaire et exoplanètes, Ecole thématique du CNRS, Goutelas (Loire), 23 - 27 mai 2005 Edited by J.-L. Halbwachs, D. Egret, and J.-M. Hameury. Strasbourg: Observatoire astronomique de Strasbourg et Société Française d'Astronomie et d'Astrophysique (SF2A), 2006, p. 55-79, 28, 55 10

Nutzman, P. & Charbonneau, D. 2008, Publ.. Astron. Soc. Pac, 120, 317 3

Pollacco, D. L., Skillen, I., Cameron, A. C., Christian, D. J., Hellier, C., Irwin, J., Lister, T. A., Street, R. A., West, R. G., Anderson, D., Clarkson, W. I., Deeg, H., Enoch, B., Evans, A., Fitzsimmons, A., Haswell, C. A., Hodgkin, S., Horne, K., Kane, S. R., Keenan, F. P., Maxted, P. F. L., Norton, A. J., Osborne, J., Parley, N. R., Ryans, R. S. I., Smalley, B., Wheatley, P. J., & Wilson, D. M. 2006, Publ.. Astron. Soc. Pac, 118, 1407 6

Pont, F., Bouchy, F., Queloz, D., Santos, N. C., Melo, C., Mayor, M., & Udry, S. 2004, Astron. Astrophys, 426, L15 39

Pont, F., Tamuz, O., Udalski, A., Mazeh, T., Bouchy, F., Melo, C., Naef, D., Santos, N. C., Moutou, C., Diaz, R. F., Gieren, W., Gillon, M., Hoyer, S., Kubiak, M., Mayor, M., Minniti, D., Pietrzynski, G., Queloz, D., Ramirez, S., Ruiz, M. T., Shporer, A., Soszyński, I., Szewczyk, O., Szymański, M. K., Udry, S., Ulaczyk, K., Wyrzykowski, Ł., & Zoccali, M. 2008, Astron. Astrophys, 487, 749 39

Pont, F., Zucker, S., & Queloz, D. 2006, Mon. Not. R. Astron. Soc, 373, 231 17, 20, 21, 81

Régulo, C., Almenara, J. M., Alonso, R., Deeg, H., & Roca Cortés, T. 2007, Astron. Astrophys, 467, 1345 26, 32, 55, 58, 115, 130

Régulo, C. & Roca Cortés, T. 2002, Astron. Astrophys, 396, 745 34

Rossiter, R. A. 1924, Astrophys. J, 60, 15 12

Sackett, P. D. 1999, in NATO ASIC Proc. 532: Planets Outside the Solar System: Theory and Observations, ed. J.-M. Mariotti & D. Alloin, 189–+ 5

Samus, N. N., Durlevich, O. V., & et al. 2009, VizieR Online Data Catalog, 1, 2025 128

Scargle, J. D. 1982, Astrophys. J, 263, 835 115, 131

Scholz, A. & Eislöffel, J. 2004, Astron. Astrophys, 419, 249 105

Seager, S. & Mallén-Ornelas, G. 2003, Astrophys. J, 585, 1038 2, 6, 7, 8

Seager, S. & Sasselov, D. D. 2000, Astrophys. J, 537, 916 11

Sirko, E. & Paczyński, B. 2003, Astrophys. J, 592, 1217 11, 42, 132

Snellen, I. A. G., Koppenhoefer, J., van der Burg, R. F. J., Dreizler, S., Greiner, J., de Hoon, M. D. J., Husser, T. O., Krühler, T., Saglia, R. P., & Vuijsje, F. N. 2009, Astron. Astrophys, 497, 545 39

Swain, M. R., Vasisht, G., & Tinetti, G. 2008, Nature, 452, 329 vii, 11, 12

Tamuz, O., Mazeh, T., & Zucker, S. 2005, Mon. Not. R. Astron. Soc, 356, 1466 18, 19, 130, 131

Tingley , B. 2003, Astron. Astrophys, 408, L5 25, 31

Tingley, B. . 2004, Astron. Astrophys, 425, 1125 11

Tingley, B. 2003, Astron. Astrophys, 403, 329 25

Tingley, B. & Sackett, P. D. 2005, Astrophys. J, 627, 1011 11, 40, 47, 49, 132

Tody, D. 1986, in Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, Vol. 627, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. D. L. Crawford, 733–+ 97

Udalski, A., Paczynski, B., Zebrun, K., Szymanski, M., Kubiak, M., Soszynski, I., Szewczyk, O., Wyrzykowski, L., & Pietrzynski, G. 2002a, Acta Astronomica, 52, 1 6, 18, 21, 31, 39, 40, 97, 129, 130

Udalski, A., Pietrzynski, G., Szymanski, M., Kubiak, M., Zebrun, K., Soszynski, I., Szewczyk, O., & Wyrzykowski, L. 2003, Acta Astronomica, 53, 133 21, 31, 39, 41, 97, 129, 130

Udalski, A., Pont, F., Naef, D., Melo, C., Bouchy, F., Santos, N. C., Moutou, C., Díaz, R. F., Gieren, W., Gillon, M., Hoyer, S., Mayor, M., Mazeh, T., Minniti, D., Pietrzyński, G., Queloz, D., Ramirez, S., Ruiz, M. T., Shporer, A., Tamuz, O., Udry, S., Zoccali, M., Kubiak, M., Szymański, M. K., Soszyński, I., Szewczyk, O., Ulaczyk, K., & Wyrzykowski, Ł. 2008, Astron. Astrophys, 482, 299 39

Udalski, A., Szewczyk, O., Zebrun, K., Pietrzynski, G., Szymanski, M., Kubiak, M., Soszynski, I., & Wyrzykowski, L. 2002b, Acta Astronomica, 52, 317 21, 31, 39, 41, 97, 129, 130

Udalski, A., Szymanski, M. K., Kubiak, M., Pietrzynski, G., Soszynski, I., Zebrun, K., Szewczyk, O., & Wyrzykowski, L. 2004, Acta Astronomica, 54, 313 21, 31, 39, 41, 97, 129, 130

Udalski, A., Zebrun, K., Szymanski, M., Kubiak, M., Soszynski, I., Szewczyk, O., Wyrzykowski, L., & Pietrzynski, G. 2002c, Acta Astronomica, 52, 115 21, 31, 39, 40, 97, 129, 130

Winn, J. N., Noyes, R. W., Holman, M. J., Charbonneau, D., Ohta, Y., Taruya, A., Suto, Y., Narita, N., Turner, E. L., Johnson, J. A., Marcy, G. W., Butler, R. P., & Vogt, S. S. 2005, Astrophys. J, 631, 1215 vii, 12, 13, 14

Wolszczan, A. & Frail, D. A. 1992, Nature, 355, 145 2

Acknowledgements

This time of my life has been a great learning experience in all senses. First of all, I would like to thank my supervisor Prof. Thomas Henning, for all his support during these years at the Max Planck Institute. I appreciate his accurate advice when I needed it.

I would like to thank Prof. Reinhard Mundt for all his advices during the last stages of my PhD. I am also grateful of my colleague and friend Nikolay Nikolov, I thank him for all his support.

I appreciate the help of Christian Fendt, who was always available to solve all kind of possible problems.

I would like to thank Prof. Joachim Wambsganss and Prof. Bernd Jaehne for kindly agreeing to form part of my examination committee.

I thank my former supervisors Edgardo Costa and Rene Mendez. Thanks to them I had the opportunity to come here.

I thank all my friends that helped and encouraged me in really difficult moments. I would like to thank Anita, Joe, Carolina, Ulrik, Paolita, Leonard, Sonita, Dading, Crystalita, Tatiana, Natashita for all those beautiful moments, I will keep you in my heart forever.

My time in Heidelberg would have not been the same without dancing, my greatest passion. Thus, I would like to thank Natashita and Tatiana for all those beautiful moments in the dance floor. Those moments will be in my memory forever.

Last but not least, Me gustaria agradecer a mi familia, a mis hermanas Eleonora y Rossana por todo el carino y apoyo durante momentos muy dificiles de mi vida aca. A mi padre por todo el apoyo y buenos consejos durante mi tiempo aca. La energia que me entregaron fue escencial para salir adelante. Finalmente a el gran Padre alla

arriba, que le dio fuerza a Max Gladiador para afrontar como el luchador que es esta gran batalla.

I would like to finish with a quote of a salsa song: “.... no hay que llorar, que la vida es un carnaval, es mas bello vivir cantando .....y las penas se van cantando... “ ... thanks for reading ;).

Declaration

I herewith declare that I have produced this paper without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. This paper has not previously been presented in identical or similar form to any other German or foreign examination board.

The thesis work was conducted from 2007 to 2010 under the supervision of Dr. Prof. Thomas Henning at Max-Planck-Institute for Astronomy (MPIA).

Heidelberg, November 08^{th}, 2010.