Searching for Exoplanets Using the Transit Method

by

Zahra Essack

A dissertation submitted in fulfilment of the academic requirements for the

degree of

Master of Science

in the School of Mathematics, Statistics and Computer Science

College of Agriculture, Engineering and Science

University of KwaZulu-Natal

Durban

South Africa

August 2017

Declaration: Supervisor

As the candidate's Supervisor, I agree to the submission of this dissertation.

Signed:

Dr. Matthew Hilton

Date: 07/08/2017

Abstract

We present a study designed to detect transiting exoplanets in Kepler light curve data. We developed an exoplanet detection algorithm based on modelling transit light curves and fitting the models to light curve data using a chi-square minimization approach in order to identify exoplanets and estimate their properties such as orbital period, planetary radius and semi-major axis (orbital radius) from the best fit parameters of the model. We applied our algorithm to a blind sample of Kepler mission data consisting of approximately 4500 stars. The selection criteria for the blind sample were $T_{star} < 6000$ K, $R_{star} < 1 R_{\odot}$ and $13.5 <$ Kepler Magnitude < 14. The blind sample contained 70 known exoplanets. Our algorithm detected 50 of the 70 known exoplanets in the blind sample. We found that our algorithm was effective in detecting exoplanets with planet-star radius ratios greater than 0.01 ($k > 0.01$) and/or exoplanets with radii greater than $2.5 R_{\oplus}$, as well as short-period exoplanets ($p < 90$ days). Twenty four of the exoplanets in the blind sample were from multi-planetary systems and, in these cases, we found our algorithm first fits for the largest transit depth and/or (subsequently) for the shortest orbital period. We did not find any potentially habitable exoplanets in our blind sample. This is not unexpected as, of more than 3400 exoplanets found to date after surveying upward of 500 000 stars, only 52

exoplanets are considered potentially habitable to varying degrees i.e. 1.5% of all exoplanets found to date are considered potentially habitable.

Declaration: Plagiarism

I, Zahra Essack, declare that

1. The research reported in this dissertation represents my own original research, except where due reference is made.

2. This dissertation has not been submitted in the past for any degree at any other institution (academic or otherwise).

3. This dissertation does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from the said persons.

4. This dissertation does not contain other persons' writing, unless specifically acknowledged as being sourced from the said persons. Where other written sources have been quoted, then:

   a. Their words have been re-written but the general information attributed to them has been referenced.

   b. Where their exact words have been used, their writing has then been placed in italics, within quotation marks, and referenced.

5. This dissertation does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the dissertation and in the Bibliography.

Signed: *[signature]*   Name: Zahra Essack   Date: August 2017

Acknowledgements

Zahra Essack

August 2017

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Extrasolar planets (or exoplanets) are planets outside our solar system orbiting stars other than our Sun (Ridpath [2012]). The first exoplanet was found in 1992, orbiting a pulsar (Wolszczan and Frail [1992]), followed by the discovery of the first exoplanet orbiting a Sun-like star in 1995 (Mayor and Queloz [1995]). Since then, more than 3400 extrasolar planets have been found and confirmed to date with over 4400 planet candidates yet to be verified[1]. More than 2300 of the confirmed exoplanets have been found by NASA's successful Kepler mission (now repurposed into a mission called K2)[1]. Extrasolar planets are a rapidly growing field in astronomy centred on developing new methods and improving existing methods to detect habitable exoplanets and find an exo-Earth in the search for life beyond Earth.

In this Chapter, we first define the different types of exoplanets that have

---

[1]http://exoplanetarchive.ipac.caltech.edu/docs/counts_detail.html

been discussed in recent works in the literature. We then describe the primary methods used to detect exoplanets, and finally we discuss factors affecting the habitability of exoplanets and methods to ascertain their habitability. We end this introduction with a description of the structure of this dissertation.

1.1 Exoplanet Categories

We begin with an introduction to the types of exoplanets currently discussed in the literature and explain some exoplanet terminology.

Giant planet Giant planets are planets that do not have a solid or liquid surface and are surrounded by an envelope of hydrogen and helium gas. Gas giants refer to planets like Jupiter and Saturn and ice giants refer to planets like Uranus and Neptune (Dotson et al. [2010]).

Terrestrial planet Terrestrial planets, also called "rocky" planets, have surfaces defined by the radius of a solid or liquid interior (core). Terrestrial planets have masses less than $10M_{\oplus}$ (Earth masses) as it is thought that planets larger than this will form giant planets (Yaqoob [2011]).

Potentially habitable planet A planet whose orbit lies within the habitable zone of its star (the region around the star in which liquid water can exist on a planet's surface). Potentially habitable planets have solid or liquid surfaces. The term "potentially habitable" implies that

some planets that fall within the habitable zone may not be habitable
- as is the case with Venus and Mars in our solar system (Dotson et al.
[2010]).

Earth-like planet/Earth Analog A planet with a mass and radius of
$1M_\oplus$ and $1R_\oplus$ (Earth radius) respectively, in an Earth-like orbit around
a Sun-like star. Earth-like planets are not necessarily habitable planets,
nor are habitable planets necessarily Earth-like. To refer to a planet
with liquid water oceans and continental land masses, the term "Earth
twin" is used (Dotson et al. [2010]).

Hot Jupiter Hot Jupiters are planets that have masses comparable to or
greater than the mass of Jupiter (but less than $13M_{Jup}$). They orbit
their host stars at very short distances e.g. 0.05 AU (astronomical units
- 1AU $\sim$ 150,000,000 km) . Some hot Jupiters may have densities less
than Saturn (i.e. 687 kg·m^{-3}) due to inflation of their atmospheres
because of their close proximity to their host stars (Yaqoob [2011]).

Super Earth Super Earths are planets that are larger than our Earth but
usually have masses $< 10M_\oplus$ and/or radii $< \pm 1.75R_\oplus$. Super Earths
refer to primarily rocky planets so the density of the planet is taken into
account in this categorization but not necessarily Earth-like properties
(Dotson et al. [2010]).

Exo-Neptune Exo-Neptunes describe exoplanets with masses between $10M_\oplus$
and $25M_\oplus$ and volumes dominated by an envelope of hydrogen/helium

gas. Their masses, however, are dominated by heavier elements (Dotson et al. [2010]). Exo-Neptunes may be further categorized as "warm" or "hot", depending on their orbital distance from their host star.

1.2   Methods of Detection

1.2.1   The Radial-Velocity Method

The first exoplanet orbiting a Sun-like star was discovered using the radial-velocity method (Mayor and Queloz [1995]). In a planet-star system, the planet and star orbit the common centre of mass of the system. The motion of a star can be resolved into three components: one component of the star's motion along the observer's line of sight, and two components of the star's motion on the plane of the sky (Dotson et al. [2010]). The radial-velocity measured by this method is along the observer's line of sight. The radial-velocity of the star can be measured using Doppler spectroscopy - observing shifts in the stellar spectrum due to the gravitational force exerted on the star by the planet, causing the star to wobble. The observed wavelength shift of a selected stellar spectral line can be expressed by the non-relativistic form of the Doppler equation

$$\frac{\Delta\lambda}{\lambda} = \frac{v}{c} \; ,$$ (1.1)
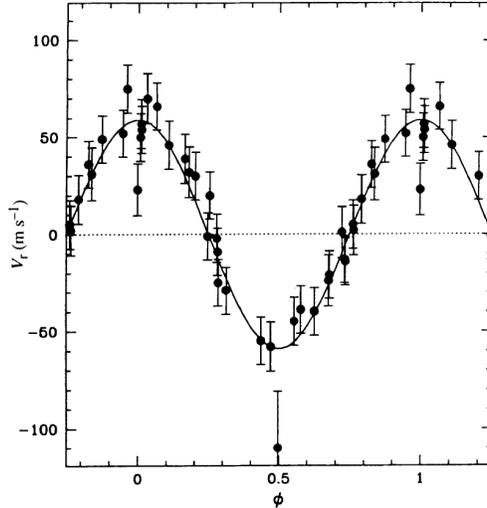
Figure 1.1: Periodic radial-velocity signal of 51 Pegasi b, the first exoplanet discovered orbiting a Sun-like star (Mayor and Queloz [1995]).

where c is the speed of light, λ is the rest frame wavelength, $\Delta\lambda$ is the observed wavelength shift, and v is the radial velocity.

The star is said to be blue-shifted if it is moving towards the Earth and red-shifted if it is moving away from the Earth. Blue-shifted implies a shorter observed wavelength while red-shifted implies a longer observed wavelength. With respect to equation (1.1) above, the radial velocity of a blue-shifted star is calculated to be negative while the radial velocity of a red-shifted star is calculated to be positive.

A periodically shifting stellar spectrum implies that the star is most probably being orbited by another body. The amplitude of radial-velocity variations seen in Figure 1.1 depends on the planet mass and orbital distance.

Using the radial-velocity method, planetary orbital period, orbital eccentricity and minimum mass can be measured. A significant drawback of the radial-velocity method is the lack of an upper bound on the mass of the body orbiting the star because we are only able to measure the radial velocity along the line of sight (Santos [2008]).

At the time of the discovery of 51 Pegasi b, a Jupiter sized exoplanet, in 1995, the radial velocity precision achieved by instruments was 15 m.s^{-1} (in contrast, Jupiter induces a 12 m.s^{-1} radial-velocity signal on the Sun). Since then, the radial velocity precision of instruments has improved to 1 m.s^{-1} (Lovis et al. [2006]). This allowed for detections of several Neptune and super Earth sized exoplanets, as seen in Figure 1.2. However, in order to detect Earth-like planets orbiting at 1 AU around Sun-like stars, the radial velocity precision required is at least 0.1 m.s^{-1} (Mayor et al. [2014]).

The radial-velocity method is used as a follow-up method to confirm planetary candidates found using the transit method by space telescopes like Kepler. Radial-velocity surveys using spectrographs such as HARPS (High Accuracy Radial velocity Planet Searcher), HARPS-N (High Accuracy Radial velocity Planet Searcher for the Northern hemisphere) and Keck HIRES (High Resolution Echelle Spectrometer) will explore planetary systems in the solar neighbourhood, and will carry out the follow-up of the transit search missions to measure planet masses (Mayor et al. [2014]).
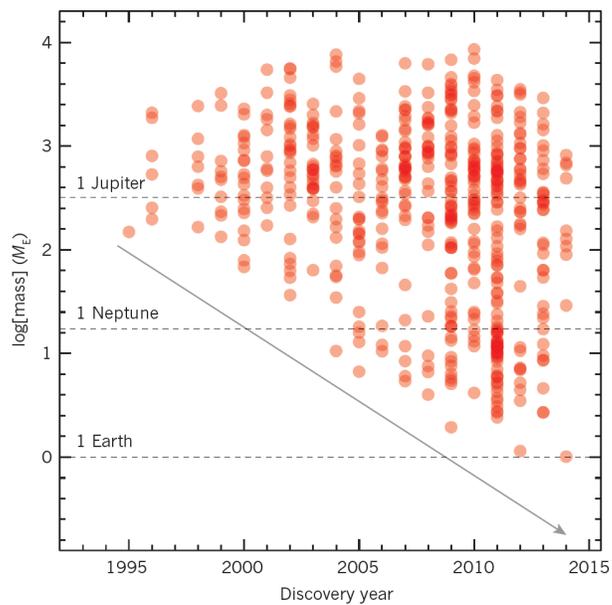
Figure 1.2: Plot of minimum mass of exoplanet as a fuction of discovery year, for exoplanets discovered using the radial-velocity method. This plot illustrates the decrease in mass of the discovered exoplanets as the precision of radial-velocity surveys increases. (Mayor et al. [2014]).
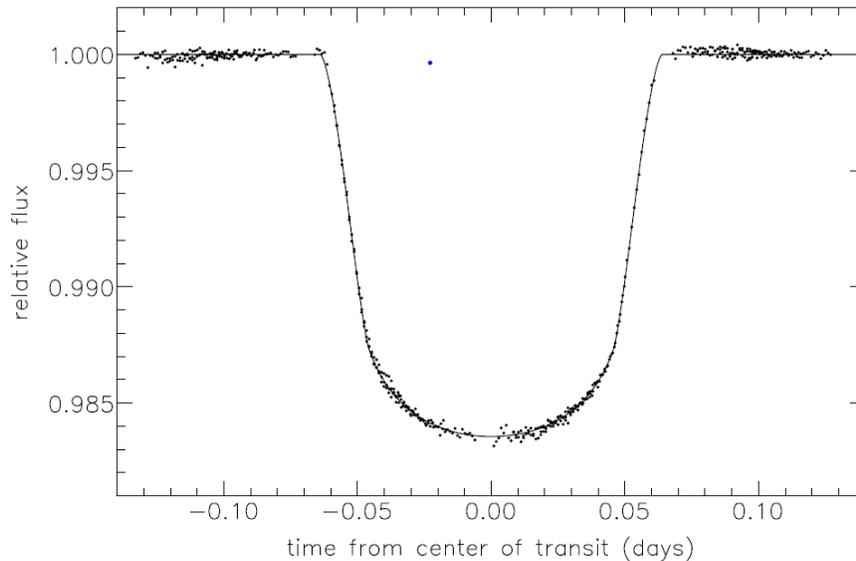
Figure 1.3: Transit of an exoplanet (HD 209458 b) across the face of its host star (Brown et al. [2001]).

1.2.2 The Transit Method

A transit occurs when an exoplanet passes across the face of its host star and blocks out some of the star's light. Upon studying the light curve of the star, a plot of brightness (flux) vs time, a dip in the brightness of the star can be observed (Figure 1.3). If these dips in brightness occur periodically, it can be indicative of the possible presence of an exoplanet orbiting the star.

In order for an exoplanet transit to be observed, the exoplanet must have an edge-on orbit around its star, i.e. the inclination of the exoplanet orbit must be close to 90° as observed from Earth.

The change in the brightness of the star is related to the radius of the

orbiting planet by the equation

$$\frac{\Delta L}{L} = \left(\frac{R_{planet}}{R_{star}}\right)^2 , \tag{1.2}$$

where ΔL is the change in brightness, L is the star's brightness and R_{planet} and R_{star} are the radii of the transiting exoplanet and star respectively. $\frac{R_{planet}}{R_{star}}$ is called the planet-star radius ratio and is represented by k. For a Jupiter-sized planet, $k \approx 0.1$, leading to a transit depth of approximately 1%. This is approximately 0.5% less than the transit depth seen in Figure 1.3, however exoplanet HD 209458 b has a radius of approximately $1.380 R_{Jup}$ (Jupiter radii). For an Earth-sized planet, $k \approx 0.01$ (Santos [2008]).

Transit light curves can be used to measure the planetary orbital period, the planet-star radius ratio and the orientation of the planet's orbit. Transit timing variations (TTVs) occur when an exoplanet's orbit is perturbed by gravitational interactions with additional exoplanets in the stellar system. TTVs can be used to infer the presence of unseen bodies within the stellar system by their interaction with transiting exoplanets. If two or more exoplanets transit the same star it is possible to precisely determine the masses of the exoplanets and hence calculate their densities to determine their composition (Holman and Murray [2005]).

Kepler's Third Law

The planetary orbital period measured from the transit light curve can be used to calculate the semi-major axis of the exoplanet's orbit using Kepler's third law, assuming circular orbits and that the mass of the planet is negligible.

Johannes Kepler deduced three laws of planetary motion based on analyzing Tycho Brahe's observations of the planets in our solar system. Kepler published his first two laws in 1609 and his third law of planetary motion in 1619 (Russell [1964]). Kepler's third law states:

"The square of a planet's orbital period about the Sun (in years) is equal to the cube of its semi-major axis."

This version of Kepler's third law only worked for planets orbiting the Sun and Kepler could not explain why his laws worked. Isaac Newton was able to derive all of Kepler's laws from his three laws of motion and his law of universal gravitation in 1697 (Russell [1964]). Kepler's third law was then generalized to apply to objects not orbiting the Sun and is given by

$$P^2 = \frac{4\pi^2 a^3}{G(M_{star} + M_{planet})} \ , \tag{1.3}$$

where P is the orbital period, a is the semi-major axis of the orbit, G is universal gravitational constant and M_{star} and M_{planet} are the masses of the star and planet respectively.

If $M_{star} \gg M_{planet}$, the approximation $M_{star} + M_{planet} \approx M_{star}$ can be made. If P is in units of years, a in units of AU and M_{star} in units of $M_\odot$ (solar masses), then equation (1.3) can be written as[2]

$$\frac{P^2}{yrs} = \left(\frac{a^3}{AU}\right)\left(\frac{M_\odot}{M_{star}}\right) \ , \tag{1.4}$$

For a given system, the probability that a full transit will occur is expressed by,

$$p = \frac{R_{star}}{a} \ , \tag{1.5}$$

where R_{star} is the stellar radius and a is the planetary orbital radius (semi-major axis), assuming circular orbits (Santos [2008]).

From equation (1.5), the probability of observing transits of Earth-like planets is around 0.5%. This shows that transits of short-period planets have a higher probability of being detected than transits of long-period planets i.e. planets with semi-major axis lengths comparable to 1 AU. The transit method is biased toward detecting large, short period planets as opposed to Earth-like planets with long orbital periods (Dotson et al. [2010]). Transit surveys require a sufficiently large number of stars to be observed to increase the probability of observing transits, since the 90° alignment with Earth required has a very low likelihood, in order to obtain statistically significant

---

[2]exoplanetarchive.ipac.caltech.edu/docs/poet_calculations.html

results (Borucki et al. [2010]).

Transit method observations are followed up with radial-velocity measurements to estimate the mass of the transiting object, in order to classify the object as an exoplanet. 80% of the total number of exoplanets discovered to date have been found using the transit method.

There are a number of currently active exoplanet transit surveys. Space based surveys include NASA's Kepler space telescope - now in mission phase K2 (Borucki et al. [2010], Howell et al. [2014]) - and the European Space Agency's spaced based CoRoT (Convection, Rotation and planetary Transits) telescope (Baglin et al. [2002]) .

There are several ground based transit surveys including KELT (Kilodegree Extremely Little Telescope) which consists of KELT-North in Arizona in the United States (Pepper et al. [2007]), and KELT-South at the South African Astronomical Observatory (SAAO) in Sutherland (Pepper et al. [2012]), HATNet - Hungarian-made Automated Telescope Network (Bakos et al. [2004]) and SuperWASP (Super Wide Angle Search for Planets) which consists of wide-field imaging cameras at the Observatorio del Roque de los Muchachos on La Palma in the Canary Islands, and the Sutherland Station of the SAAO (Pollacco et al. [2006]).

In 2018, NASA's TESS (Transiting Exoplanet Survey Satellite) mission will be launched. TESS's goal is to detect small planets with bright host stars in the solar neighborhood in an all sky survey of 200,000 stars, in order
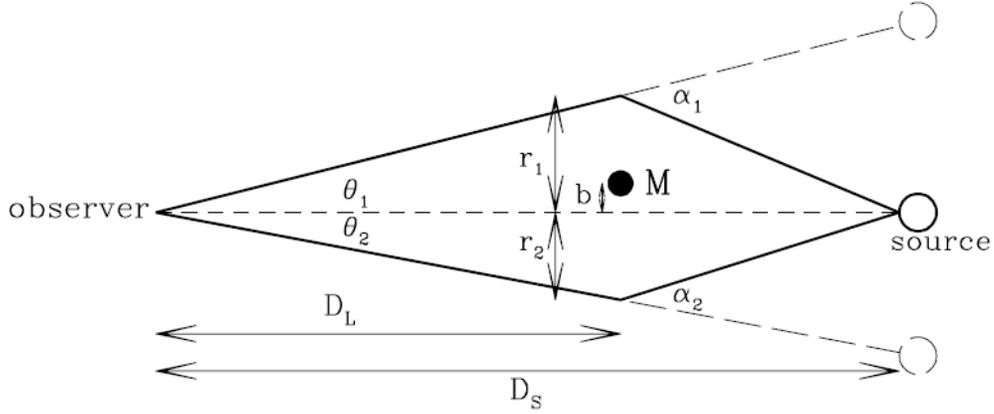
Figure 1.4: Two images, offset by angles θ_1 and θ_2, of the source are observed due to gravitational lensing of mass M (Bennett [2008]).

to characterize the exoplanets and their atmospheres[3].

1.2.3   The Gravitational Microlensing Method

In Albert Einstein's theory of General Relativity, gravity bends the path of light. As a result, if a massive object lies between an observer and a target, the observer will see two images offset by angles from the line of sight of the target, as seen in Figure 1.4 (Bennett [2008]). In general, gravitational lensing can produce more than two images depending on the distribution of matter in the lens and the source-lens-observer geometry.

From General Relativity, gravitational lensing requires the deflection angle α given by

---

[3]https://tess.gsfc.nasa.gov/overview.html

13

$$\alpha = \frac{4GM}{c^2 r} \ , \tag{1.6}$$

where G is the universal gravitational constant, M is the mass of the object the light ray is traveling past, c is the speed of light and r is the impact parameter (Bennett [2008]).

Gravitational microlensing occurs when light from a distant, bright source star is bent around a lensing object (usually a star) close to the observer's line of sight. The lensing star magnifies the source star. If the lensing star has planets orbiting it, spikes (or planetary deviations - see Figure 1.5) can be observed in the resulting microlensing light curve of the lensing star. This spike is a signature of an exoplanet. The planet mass and planet-star separation can be obtained (Bond et al. [2004]).

Gravitational microlensing is a technique that is sensitive to small exoplanets and large planet-star separations (1.5 - 4 AU). This method would allow for the discovery of Earth-sized planets possibly in the habitable zones of their stars. While the radial velocity and transit methods aid the search for large, short-period exoplanets, gravitational microlensing aids the search for small, long-period exoplanets. Unlike the transit method, gravitational microlensing does not depend upon light from the host star to detect planets. This could lead to discoveries of exoplanets around unseen stars (Bennett [2008]).

Gravitational microlensing occurrences are rare and depend on the chance
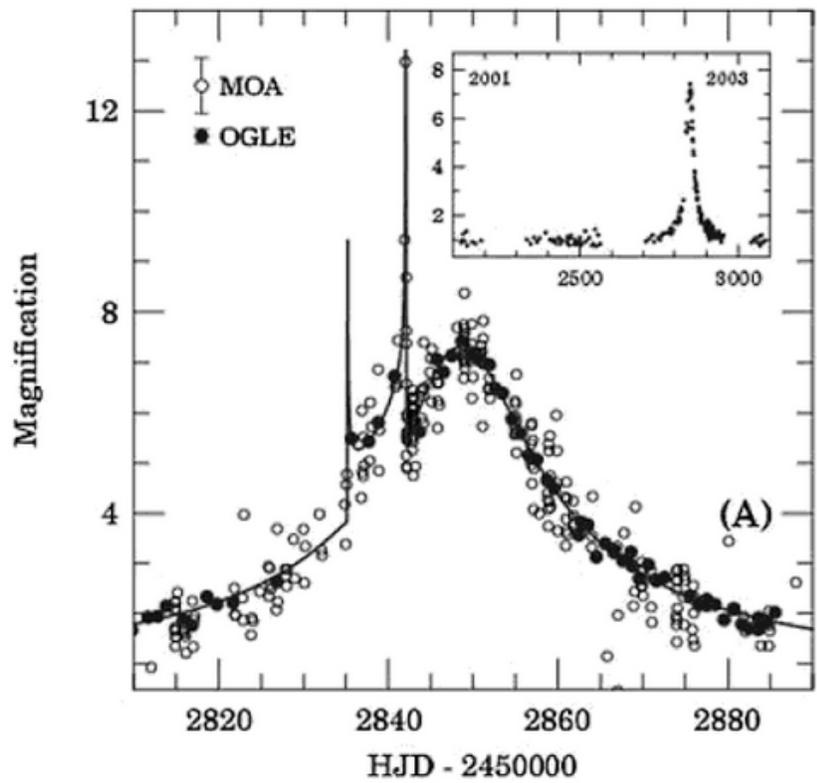
Figure 1.5: Microlensing light curve of a planet observed by OGLE (Optical Gravitational Lensing Experiment) (Bond et al. [2004]).

alignment of lensing and source stars. The microlensing event only occurs once, hence no follow up observations or measurements can be made (Dotson et al. [2010]). Only 46 out of more than 3400 confirmed exoplanets have been discovered using the gravitational microlensing method (1.3%) [4]. While this method is effective in the search for potentially habitable Earth-sized planets, it would require a constant all sky survey to detect these chance microlensing events. This is possible in the near future with the LSST (Large Synoptic Survey Telescope), designed to survey the entire night sky twice each week[5].

1.2.4 Other Methods

Other methods of detection include astrometry (the reflex motion of a star caused by an orbiting planet can be observed using high accuracy observations of the star's position on the sky), direct imaging of exoplanets (Figure 1.6 - difficult due to the brightness of the star it orbits), and pulsar timing variations (Mason [2008]). Forty-four exoplanets have been detected using direct imaging, six using pulsar timing variations, and just one using astrometry.

---

[4]https://exoplanetarchive.ipac.caltech.edu/docs/counts_detail.html
[5]https://www.lsst.org/lsst/
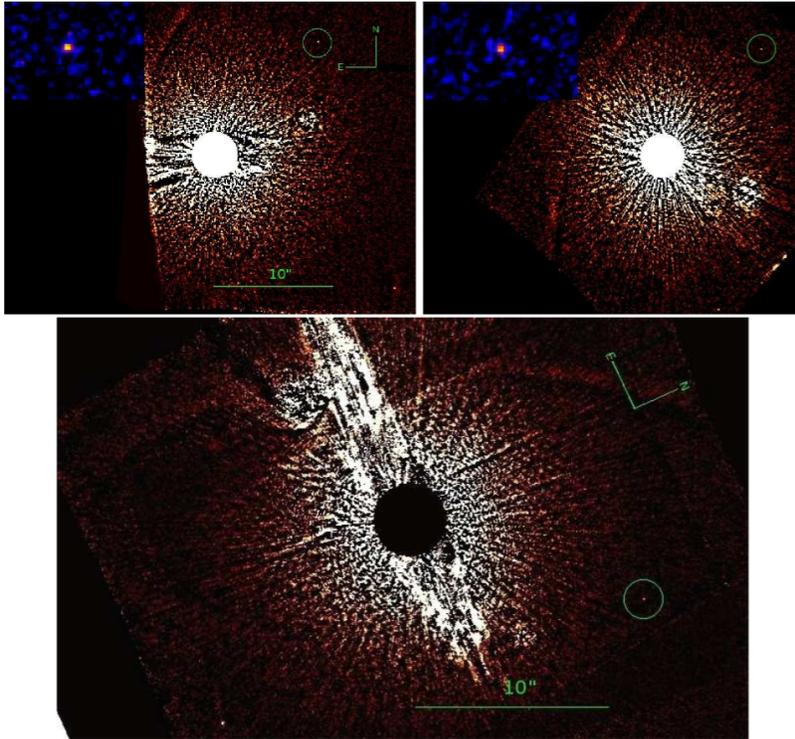
Figure 1.6: Direct imaging of star Fomalhaut from 2006 centred on its planet, Fomalhaut b, using HST/ACS (Hubble Space Telescope Advanced Camera for Surveys) data (Currie et al. [2012]).

1.3 Exoplanet Habitability

1.3.1 The Habitable Zone

The habitable zone (HZ) is the region around a star in which rocky planets can maintain liquid water on their surfaces (Kasting et al. [2014]). The continuous habitable zone (CHZ) is the region around a star that remains habitable over some finite period of time. As stars age, their luminosity increases, resulting in the star's habitable zone moving further outward (Dotson et al. [2010]).

The inner and outer boundaries of the habitable zone for our solar system are based on the 1-D, cloud-free, climate model calculations in Kasting et al. [1993], assuming both Venus and Mars were habitable planets in the past. The habitable zone boundaries are scaled for solar luminosity ($L_\odot$) and it is assumed that planet albedo (the fraction of total radiation reflected from the planet (Ridpath [2012])) is constant with wavelength. The inner radius of the habitable zone is given by

$$R_{HZ,inner} = 0.75\sqrt{\frac{L_{star}}{L_\odot}} \ , \tag{1.7}$$

and the outer radius of the habitable zone is given by

$$R_{HZ,outer} = 1.77\sqrt{\frac{L_{star}}{L_\odot}} \ , \tag{1.8}$$

where L_{star} is the star's luminosity and the constants 0.75 and 1.77 are the

(optimistic) inner and outer boundaries of our solar system's habitable zone. $R_{HZ,inner}$ and $R_{HZ,outer}$ are in units of AU[6]. Equations (1.7) and (1.8) are used to calculated habitable zone boundaries in this dissertation.

A more conservative approach to estimating habitable zone boundaries is described in Kopparapu et al. [2013], taking into account the absorption coefficients of water and carbon dioxide. The inner habitable zone boundary is now determined by loss of water via the runaway greenhouse effect (e.g. Venus) and the outer habitable zone boundary is determined by carbon dioxide condensation (e.g. Mars), as shown in Figure 1.7 (Seager [2013]).

The potential habitability of an exoplanet can be inferred by its position relative to the habitable zone range of its star. However, there are several other factors that determine the habitability of a planet such as stellar irradiation, atmospheric composition (especially greenhouse gases), plate tectonics, surface albedo etc. (Seager [2013]).

1.3.2 Biosignature Gases

Biosignature gases are gases that are produced by life and accumulate to detectable levels in the exoplanet atmosphere. This allows for their possible detection by telescopes using atmospheric spectroscopy. Only globally mixed, spectroscopically active biosignature gases will be visible in an exoplanet spectrum from afar (Seager [2013]).

---

[6]https://exoplanetarchive.ipac.caltech.edu/docs/poet_calculations.html
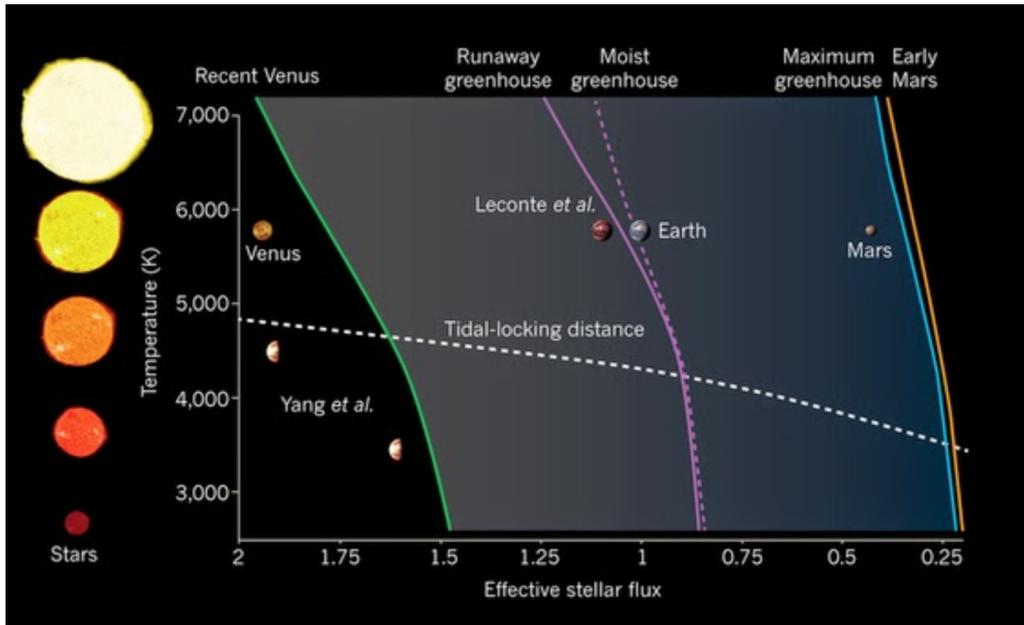
Figure 1.7: Location of optimistic and conservative habitable zone boundaries for different stellar temperatures and fluxes (adapted from Yang et al. [2014]).

Requirement	Occurrence in the Solar System
Energy	Common
Predominately light	Photosynthesis at 100 AU light levels
Chemical energy	e.g., $H_2 + CO_2 \rightarrow CH_4 + H_2O$
Carbon	Common as CO_2 and CH_4
Liquid water	Rare, only on Earth for certain
N,P, S, Na, and other elements	Likely to be common

Figure 1.8: Ecological requirements for life (McKay [2014]).

Requirement	Note
1. Temperature and state of water	T between -15 °C and 122 °C, $P > \sim 0.01$ atmospheres
2. Water availability	Few days per y of rain, fog, snow, or RH > 80%
3. Light and redox energy sources	
4. UV and Ionizing radiation	Limits exemplified by *D. radiodurans* (Table 3)
5. Nitrogen	Enough N_2 for fixation or fixed nitrogen present
6. O_2	Over 0.01 atmospheres need to support complex life

Figure 1.9: Checklist of requirements for exoplanet habitability (McKay [2014]).

In the search for extraterrestrial life, the underlying assumption is that extraterrestrial life has the same ecological and environmental requirements as life on Earth (see Figures 1.8 and 1.9) and goes through the same chemical reactions and metabolic processes that produce biosignature gases on Earth (McKay [2014], Seager [2013]).

The dominant biosignature gases found on Earth are N_2O (nitrous oxide), O_2 (molecular oxygen) and its photochemical product O_3 (ozone), and

possibly CH_4 (methane) on early Earth (Seager [2013]). Other biosignature gases considered include CO_2 (carbon dioxide), SO_2 (sulfur dioxide) and H_2S (hydrogen sulfide). However, many biosignature gases can be interpreted as "false positives" because they can be produced abiotically. For example, H_2S and SO_2 are produced by volcanism and hence are not a definitive sign of life (Seager and Bains [2015]).

Biosignature gases are divided into categories based on their likelihood to be a definitive indicator for extraterrestrial life.

Category 1 Biosignature gases that are by-products of metabolic reactions that involve energy capture from environmental redox (reduction and oxidation) reactions (chemical potential gradients) e.g. CH_4 produced from methanogenesis (formation of methane by microbes known as methanogens). These gases are abundant because they can be formed from plentiful chemicals in the environment but have a high probability of being false positives because they can be created by geological process, not exclusively by life (Seager [2014]).

Category 2 Biosignature gases produced by life for reasons other than energy capture. These chemicals have a wider variety and are produced in much smaller quantities. They are expected to have a lower probability of being false positives (Seager [2014]).

Category 3 Biosignature gases that are highly specialized chemicals, produced for organism-specific reasons e.g. isoprene (organic compound

produced by plants) and methyl chloride (produced by polar macroalgae). Extraterrestrial organisms are assumed to be similar to organisms on Earth in order to produce the examples of organism-specific chemicals listed (Seager [2014]).

1.3.3 Ascertaining Habitability

In order to ascertain an exoplanet's habitability and, further, detect extraterrestrial life, its atmosphere must be remotely observed in order to obtain atmospheric spectra of chemicals and, potentially, biosignature gases. There are currently two exoplanet atmosphere observation methods.

Direct Imaging

Direct imaging involves taking pictures of potentially habitable exoplanets. The difficulty involved with directly imaging exoplanets is the sheer brightness of the planet's host star as compared to the exoplanet (host stars tend to be millions of times brighter than exoplanets). Currently, this method is limited to imaging big, bright, young or massive planets orbiting at large distances from their host stars and is made possible with large ground-based telescopes and adaptive optics - to cancel out Earth's atmospheric blurring effects (Seager and Deming [2010]).

In order to directly image small (Earth-sized) exoplanets, two techniques are under development to block out starlight and focus on the orbiting exo-

planet. Both techniques require the use of space telescopes to avoid Earth's atmospheric interference.

Technique one involves inserting a coronograph (telescopic attachment designed to block out direct starlight to resolved objects hidden in the stellar glare) into the space telescope itself. This requires a highly specialized and highly stable observation system (Trauger et al. [2010]). The James Webb Space Telescope, the successor of the Hubble space telescope, has several coronographs that have the ability to directly image exoplanets (Boccaletti et al. [2005]).

Technique two is a starshade-telescope system, designed such that the starshade is flown in formation with the telescope to cast a dark and controlled shadow to block out the star light (see Figure 1.10). The starshade is a carefully shaped screen, tens of meters in diameter, with a separation distance of tens of thousands of kilometers from the telescope. The shadow cast by the starshade blocks out the starlight and leaves only the planet's reflected light, allowing for spectroscopic observations of its atmosphere (Seager [2014]).

Transit Spectroscopy

When a planet transits its host star, the starlight will pass through the atmosphere of the planet and the atmospheric features of the planet will be embedded on the starlight (the atmosphere is the blue ring around the planet in Figure 1.11). By dividing the combined planet-star spectrum by
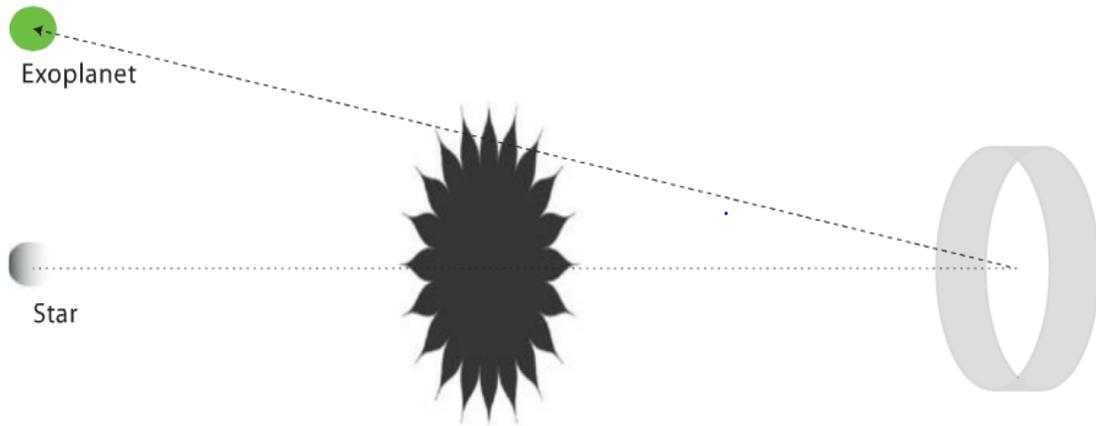
Figure 1.10: Schematic digram of starshade (middle) flying in formation with telescope (right) to block out star light in order to directly image exoplanet (Seager [2014]).

the spectrum of the star alone, taken before the planetary transit (also called primary eclipse - see Figure 1.11), the planet transmission spectrum can be obtained (Seager and Deming [2010], Schlawin et al. [2014]).

The total flux of the planet-star system is obtained just before the secondary eclipse (when the planet moves behind the star - see Figure 1.11). The total planet-star system flux is a combination of the star's flux as well as the flux from the "dayside" of the planet. When the planet moves behind the star, the total flux drops because there is no longer a planetary contribution. The flux drop is related to the size of the planet and star (see equation (1.2)) as well as the brightness of the planet and star at a given wavelength. By subtracting the total flux spectrum (combined planet-star spectrum) from the flux spectrum of the star alone, the planet flux spectrum
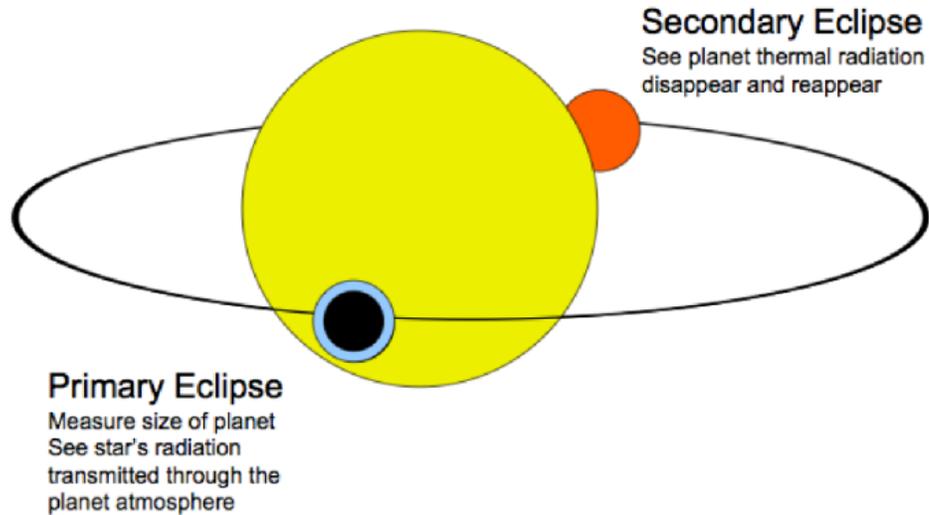
Figure 1.11: Planet transiting host star (primary eclipse) and moving behind star from observer's line of sight (secondary eclipse). Used to determine planetary transmission and flux spectrums for atmospheric spectroscopy (Seager and Deming [2010]).

can be obtained. The planet flux spectrum provides information about planetary atmospheric composition (Seager and Deming [2010]). The subtraction of spectra is incredibly difficult. After subtracting the spectra, chemical signatures are searched for in the residuals which poses the risk of systematic (repeatable) errors and false positives. If transit spectroscopy is performed from the ground, removing Earth's atmospheric interference adds a further factor of difficulty.

The transit depth, $\frac{\Delta L}{L}$, changes at different wavelengths (see Figure 1.12) because the observed planetary radius, R_{planet}, changes at different wavelengths. This is due to Rayleigh scattering - the scattering of light by parti-

cles that are much smaller than the wavelength of the light (Ridpath [2012]). Hence, the transit depth is a function of wavelength and, assuming normalized flux, is given by

$$\frac{\Delta L}{L}(\lambda) = \left(\frac{R_{planet}(\lambda)}{R_{star}}\right)^2 \ ,$$ (1.9)

Different atmospheric molecules can be observed at different wavelengths due to the different scattering cross sections of the atmosphere. The atmospheric scattering cross section as a function of wavelength is given by $\sigma(\lambda)$ (Benneke and Seager [2012]).

For the optically thick part of the atmospheric spectrum, where photons cannot pass through the atmosphere without absorption (Ridpath [2012]), the observed planetary radius changes linearly with the logarithm of the atmospheric scattering cross section i.e. $\frac{dR_{planet}(\lambda)}{d(\ln \sigma(\lambda))}$ (Benneke and Seager [2012]).

The scale height of the atmosphere, H, is the vertical distance over which the atmospheric pressure falls by a factor of e (Ridpath [2012]), and is given by

$$H = \frac{dR_{planet}(\lambda)}{d(\ln \sigma(\lambda))} = \frac{k_B T}{\mu g} \ ,$$ (1.10)

where k_B is Boltzmann's constant, T is the estimated planetary temperature, μ is the mean molecular mass of the atmosphere and g is the planet's

27
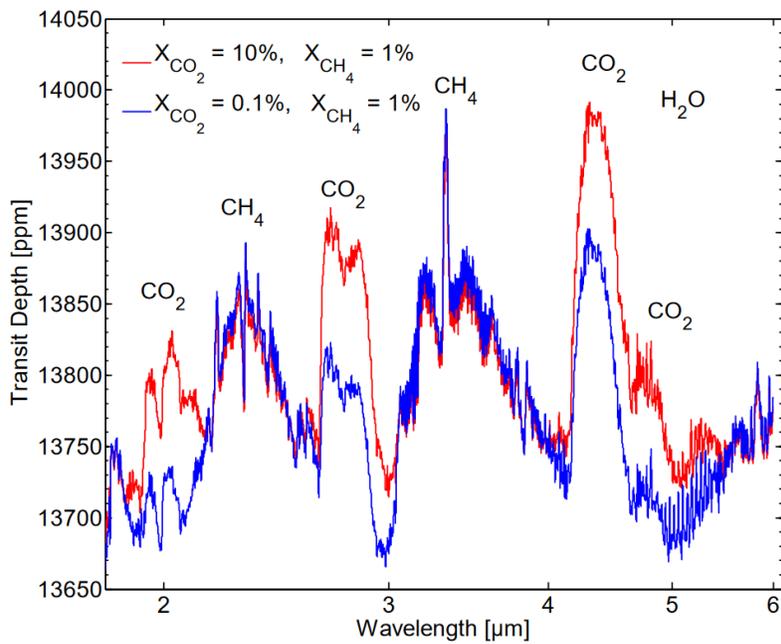
Figure 1.12: Synthetic transmission spectrum (transit depth vs wavelength) for two atmospheres with different carbon dioxide content (Benneke and Seager [2012]).

gravitational acceleration.

The mean molecular mass indicates the composition of the atmosphere. At different wavelengths, different molecules are spectroscopically active resulting in different observed planetary radii. The mean molecular mass is given by

$$\mu = \frac{k_B T}{g} \left(\frac{dR_{planet}(\lambda)}{d(\ln \sigma(\lambda))} \right)^{-1} .$$ (1.11)

It is possible to differentiate between hydrogen-dominated atmospheres (gas giant planets) and atmospheres mainly composed of water, nitrogen or carbon dioxide (terrestrial planets) because the mean molecular mass varies by a factor on the order of 8-20 (Benneke and Seager [2012]).

The atmospheres of many hot Jupiter exoplanets have been observed using transit spectroscopy (Seager and Deming [2010]) but it is not possible for Earth-sized planets orbiting Sun-like stars (G-class stars) as the planet-to-star measurement signals are too weak. Planets orbiting M-class stars, which are small, ultra cool stars, are conducive to transit spectroscopy because the planet-to-star measurement signals will be larger (Seager [2013]).

Current List of Potentially Habitable Exoplanets

Out of more than 3400 exoplanets found to date, only 52 are considered potentially habitable. The list of potentially habitable exoplanets is further divided into a conservative sample (Figure 1.13) and an optimistic sample

(Figure 1.14). The conservative sample consists of exoplanets that are more likely to have a rocky composition and maintain liquid water on their surfaces and have orbits in the conservative habitable zone described in Kopparapu et al. [2013]. Similarly, the optimistic sample consists of exoplanets that have orbits in the optimistic habitable zone (Kasting et al. [1993]) and are less likely to have a rocky composition and maintain liquid water on their surfaces.

Exoplanet type in Figures 1.13 and 1.14 is based on Planetary Habitability Laboratory's (PHL) classification of planets that includes host star spectral class (F, G, K, M), habitable zone location (hot, warm, cold) and size (miniterran, subterran, terran, superterran, Jovian, Neptunian), e.g. Earth = G-Warm Terran, Venus = G-Hot Terran, Mars = G-Warm Subterran. Of the 52 potentially habitable exoplanets, 24 were discovered during the Kepler mission [7].

February 2017 saw the exciting discovery of the TRAPPIST-1 system - a planetary system of seven planets with three planets (TRAPPIST-1 e, TRAPPIST-1 f, TRAPPIST-1 g) becoming part of the conservative sample of potentially habitable exoplanets. TRAPPIST (Transiting Planets and Planetesimals Small Telescope) is a ground-based telescope at the La Silla Observatory in Chile that employs the transit method to search for exoplanets. TRAPPIST-1 is an ultra cool M-class star located 40 light years from

---

[7]http://phl.upr.edu/projects/habitable-exoplanets-catalog

Name	Type	Mass (M$_E$)	Radius (R$_E$)	Flux (S$_E$)	T$_{eq}$ (K)	Period (days)	Distance (ly)
001. Proxima Cen b	M-Warm Terran	≥ 1.3	0.8 - 1.1 - 1.4	0.66	229	11.2	4.2
002. TRAPPIST-1 e	M-Warm Terran	0.6	0.9	0.65	229	6.1	39
003. GJ 667 C c	M-Warm Terran	≥ 3.8	1.1 - 1.5 - 2.0	0.88	247	28.1	22
004. Kepler-442 b	K-Warm Terran	8.2 - 2.3 - 1.0	1.3	0.70	233	112.3	1115
005. GJ 667 C f*	M-Warm Terran	≥ 2.7	1.0 - 1.4 - 1.8	0.56	221	39.0	22
006. Kepler-1229 b	M-Warm Terran	9.8 - 2.7 - 1.2	1.4	0.49	213	86.8	769
007. TRAPPIST-1 f	M-Warm Terran	0.7	1.0	0.38	200	9.2	39
008. LHS 1140 b (N)	M-Warm Terran	6.6	1.4	0.41	200	24.7	41
009. Kapteyn b*	M-Warm Terran	≥ 4.8	1.2 - 1.6 - 2.1	0.43	205	48.6	13
010. Kepler-62 f	K-Warm Terran	10.2 - 2.8 - 1.2	1.4	0.39	201	267.3	1200
011. Kepler-186 f	M-Warm Terran	4.7 - 1.5 - 0.6	1.2	0.29	188	129.9	561
012. GJ 667 C e*	M-Warm Terran	≥ 2.7	1.0 - 1.4 - 1.8	0.30	189	62.2	22
013. TRAPPIST-1 g	M-Warm Terran	1.3	1.1	0.26	181	12.4	39

Figure 1.13: Table of conservative sample of potentially habitable exoplanets.

Earth. The TRAPPIST-1 system is a prime candidate for atmospheric spectroscopy studies with the James Webb Space Telescope to be launched in 2018 (Gillon et al. [2017]).

1.4   Dissertation Outline

In this dissertation, we present an automated pipeline that was used to detect exoplanets in Kepler data. Chapter 2 describes the exoplanet detection algorithm we developed and describes the blind sample of Kepler data we applied it to. We present our results in Chapter 3, comparing the orbital period, planetary radius and semi-major axis values calculated by our algo-

Name	Type	Mass (M$_E$)	Radius (R$_E$)	Flux (S$_E$)	T$_{eq}$ (K)	Period (days)	Distance (ly)
001. TRAPPIST-1 d	M-Warm Subterran	0.4	0.8	1.13	263	4.0	39
002. GJ 3323 b	M-Warm Terran	≥ 2.0	0.9 - 1.3 - 1.6	1.21	264	5.4	17
003. Kepler-438 b	M-Warm Terran	4.0 - 1.3 - 0.6	1.1	1.38	276	35.2	473
004. GJ 273 b	M-Warm Terran	≥ 2.9	1.0 - 1.4 - 1.8	1.22	267	18.6	12
005. Kepler-296 e	M-Warm Terran	12.5 - 3.3 - 1.4	1.5	1.22	267	34.1	737
006. Kepler-62 e	K-Warm Superterran	18.7 - 4.5 - 1.9	1.6	1.10	261	122.4	1200
007. Kepler-452 b	G-Warm Superterran	19.8 - 4.7 - 1.9	1.6	1.11	261	384.8	1402
008. K2-72 e	M-Warm Terran	9.8 - 2.7 - 1.2	1.4	1.46	280	24.2	181
009. GJ 832 c	M-Warm Superterran	≥ 5.4	1.2 - 1.7 - 2.2	0.99	254	35.7	16
010. K2-3 d	M-Warm Terran	11.1	1.5	1.46	280	44.6	137
011. Kepler-1544 b	K-Warm Superterran	31.7 - 6.6 - 2.6	1.8	0.90	248	168.8	1138
012. Kepler-283 c	K-Warm Superterran	35.3 - 7.0 - 2.8	1.8	0.90	248	92.7	1741
013. tau Cet e*	G-Warm Terran	≥ 4.3	1.1 - 1.6 - 2.0	1.51	282	168.1	12
014. Kepler-1410 b	K-Warm Superterran	31.7 - 6.6 - 2.6	1.8	1.34	274	60.9	1196
015. GJ 180 c*	M-Warm Superterran	≥ 6.4	1.3 - 1.8 - 2.3	0.79	239	24.3	38
016. Kepler-1638 b	G-Warm Superterran	42.7 - 7.9 - 3.1	1.9	1.39	276	259.3	2866
017. Kepler-440 b	K-Warm Superterran	41.2 - 7.7 - 3.1	1.9	1.43	273	101.1	851
018. GJ 180 b*	M-Warm Superterran	≥ 8.3	1.3 - 1.9 - 2.4	1.23	268	17.4	38
019. Kepler-705 b	M-Warm Superterran	? - 12.7 - 4.8	2.1	0.83	243	56.1	818
020. HD 40307 g*	K-Warm Superterran	≥ 7.1	1.3 - 1.8 - 2.3	0.68	227	197.8	42
021. GJ 163 c	M-Warm Superterran	≥ 7.3	1.3 - 1.8 - 2.4	0.66	230	25.6	49
022. Kepler-61 b	K-Warm Superterran	? - 13.8 - 5.2	2.2	1.27	267	59.9	1063
023. K2-18 b	M-Warm Superterran	? - 16.5 - 6.0	2.2	0.92	250	32.9	111
024. Kepler-1606 b	G-Warm Superterran	? - 11.9 - 4.5	2.1	1.41	277	196.4	2869
025. Kepler-1090 b	G-Warm Superterran	? - 16.8 - 6.1	2.3	1.20	267	198.7	2289
026. Kepler-443 b	K-Warm Superterran	? - 19.5 - 7.0	2.3	0.89	247	177.7	2540
027. Kepler-22 b	G-Warm Superterran	? - 20.4 - 7.2	2.4	1.11	261	289.9	619
028. GJ 422 b*	M-Warm Superterran	≥ 9.9	1.4 - 2.0 - 2.6	0.68	231	26.2	41
029. K2-9 b	M-Warm Superterran	? - 16.8 - 6.1	2.2	1.38	276	18.4	359
030. Kepler-1552 b	K-Warm Superterran	? - 25.2 - 8.7	2.5	1.11	261	184.8	2015
031. Kepler-1540 b	K-Warm Superterran	? - 26.2 - 9.0	2.5	0.92	250	125.4	854
032. GJ 3293 d	M-Warm Superterran	≥ 7.6	1.3 - 1.9 - 2.4	0.57	220	48.1	58
033. Kepler-298 d	K-Warm Superterran	? - 26.8 - 9.1	2.5	1.29	271	77.5	1545
034. KIC-5522786 b	A-Warm Terran	5.8 - 1.8 - 0.8	1.2	2.70	305	757.2	-
035. Kepler-174 d	K-Warm Superterran	? - 14.8 - 5.5	2.2	0.43	206	247.4	1174
036. Kepler-296 f	M-Warm Superterran	28.7 - 6.1 - 2.5	1.8	0.34	194	63.3	737
037. GJ 682 c*	M-Warm Superterran	≥ 8.7	1.4 - 1.9 - 2.5	0.37	198	57.3	17
038. Wolf 1061 d	M-Warm Superterran	≥ 5.2	1.2 - 1.7 - 2.2	0.28	182	67.3	14
039. KOI-4427 b*	M-Warm Superterran	38.5 - 7.4 - 3.0	1.8	0.24	179	147.7	782

Figure 1.14: Table of optimistic sample of potentially habitable exoplanets.

rithm to the literature. In Chapter 4, we discuss our findings as compared with the literature and make some suggestions for improvement. Chapter 5 is a conclusion of our findings and suggestions for future work.

Chapter 2

Method

In this Chapter, we describe a pipeline that we developed to search for exoplanets in Kepler light curve data. We begin with a short description of the Kepler mission and the public data products that we used. We then describe the algorithm, the calibration and testing of the algorithm, and conclude with applying the algorithm to a blind sample of Kepler mission data.

2.1   The Kepler Mission

Launched in 2009, NASA's Kepler space telescope was created with the primary goal of detecting Earth-like planets orbiting Sun-like stars. Kepler surveyed a field of the Milky Way (RA=19h 22m 40s, Dec=+44° 30' 00') containing approximately 200,000 stars. The telescope was designed with a 1.4 m diameter $f/1$ primary mirror, a 0.95 m aperture and has a 115.6 deg^2

field-of-view. The aperture size is required to obtain a 4σ detection of a single transit of an Earth-sized planet transiting a 12th magnitude Sun-like star with a 6.5 hour transit duration (Borucki et al. [2010]). Kepler was able to achieve unparalleled photometric precision for 12th magnitude stars with data binned in 6.5 hour time intervals - a significant time interval since Earth takes 13 hours to transit the Sun as viewed by a distant observer (Lissauer et al. [2014]). The 6.5 hour data binning was done to reduce the total variance of the CCDs readout noise (Gilliland et al. [2011]). The 6.5 hour binning is significant with respect to Earth since 6.5 hours is half of Earth's transit time and a reduced variance in readout noise at 6.5 hour binning increases the chance of an Earth-sized planet being detected.

Kepler's four year mission ended in May 2013 after the failure of a second reaction wheel (required for precise and stable pointing away from the spacecraft's orbital plane) which occurred after the baseline mission duration of 3.5 years. Kepler's new mission is called K2, which allowed the telescope to continue being used for scientific observation.

The K2 mission involves observations of a sequence of patches of sky, called campaigns, in the ecliptic plane. This minimizes the torque exerted on the spacecraft by solar wind pressure, which reduces the pointing drift, and allows for effective control of the spacecraft despite the loss of two reaction wheels (Howell et al. [2014]). Because the K2 mission operates in the ecliptic plane, each campaign is limited to a duration of approximately 80 days to ensure the telescope is always pointing away from the Sun (Howell et al.

[2014]). The K2 mission requires the telescope to search for planets around small stars (smaller than the Sun) with orbital periods of less than one month (roughly 3 transits are required to confirm an exoplanet detection), among other tasks (Lissauer et al. [2014]).

Kepler mission targets are observed in long cadence (flux measurements obtained every 29.4 minutes), or short cadence (flux measurements obtained every 58.35 seconds). Long cadence targets are observed for three months while short cadence targets are observed for one month (Borucki et al. [2010]).

One of Kepler's pre-launch goals was to "determine the frequency of Earth-size and larger planets in or near the habitable zone of a wide variety of spectral types of stars", hence the decision to use Kepler mission data in this dissertation (Lissauer et al. [2014]). Kepler mission data is also freely available to the public.

The Kepler mission has led to the discoveries of exoplanets with diverse physical properties and characteristics, like hot Jupiters and warm Neptunes (Borucki et al. [2010]). Kepler is responsible for the discoveries of most of the known exoplanets to date (see Figure 2.1) and has been the most successful mission thus far. Kepler has made a significant contribution to the field of exoplanets and has revolutionized areas in stellar physics like astroseismology and eclipsing binary stars, but Kepler has also left us with many unanswered questions about planetary systems like the theory of planet formation, among others (Lissauer et al. [2014]).
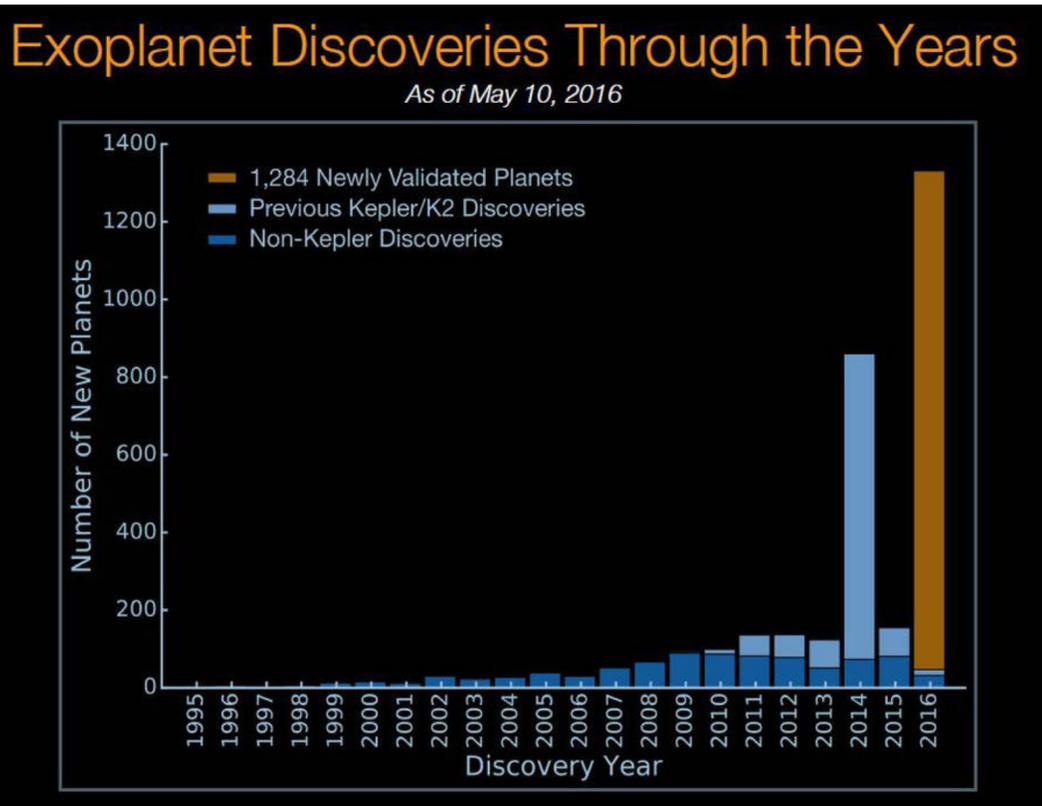
Figure 2.1: Exoplanet discoveries as of 10 May 2016 (NASA).

2.2   Algorithm

We developed an exoplanet detection algorithm to search for exoplanets using light curve data from NASA's Kepler mission. Kepler mission data was obtained from the NASA Exoplanet Archive[1]. Two quarters of long cadence light curve files (approximately 180 observation days) were used because of the frequency of the sampling - flux measurements are obtained every 29.4 minutes over 3 months - this improved our chances of exoplanet discovery. The potential habitability of an exoplanet depends on its host star and is an intrinsic property of the planet-star system. Our ability to determine an exoplanet's habitability depends on our duration of observation, given a maximum period of 3 months in this case.

Kepler light curves are associated with stars, and each star has a unique KIC (Kepler Input Catalog) identification. The light curve files contain time series data, flux data and their associated errors and corrections (Thompson et al. [2012]).

The PDCSAP_FLUX (Pre-search Data Conditioning Simple Aperture Photometry Flux) values, and their associated 1σ errors, PDCSAP_FLUX_ERR, were used from the light curve files. These flux values have been passed through the Kepler science processing pipeline to remove systematic error sources such as pointing drift, focus changes, and thermal transients (Jenkins et al. [2010]).

---

[1]http://exoplanetarchive.ipac.caltech.edu/applications/ETSS/Kepler_index.html

The PDCSAP_FLUX went through further preprocessing before it was passed through the exoplanet detection algorithm. We applied a median filter to the flux data and divided the light curve by the median filtered light curve in order to normalize the flux scale and whiten the data at the same time. A median filter can be used if the time-scale of the transits is shorter than the time-scale for any contribution from stellar variations in the light curves (Aigrain and Irwin [2004]). We set the median filter to remove all trends with a time scale longer than 1 day. The time scale of 1 day was chosen because Earth takes 13 hours to transit the Sun as viewed by a distant observer (Lissauer et al. [2014]), hence it was safe to assume a 24-hour time scale to avoid removing possible planetary transit signals from the data.

Whitening the data implies that the scatter around the mean ($\mu = 1$) should be constant and, in all likelihood, Gaussian. We use the scatter around the mean as a measurement of the 1σ flux error instead of the PDSCAP_FLUX_ERR measurements. We executed a 3σ clipping to identify and remove outliers and estimate the correct 1σ flux error on the flux data. To verify that the estimated 1σ flux error was correct, we plotted the normalized flux data as a histogram and over plotted a Gaussian with $\mu = 1$ and σ equal to the estimated 1σ flux error (see Figure 2.2).

After obtaining the flux error from the 3σ clipping and plotting the associated Gaussian distribution, we performed a skewness test to test for asymmetrical flux distributions. This was a sanity check to ensure that the 3σ clipping was executed correctly. If the distribution was skewed i.e. asymmet-
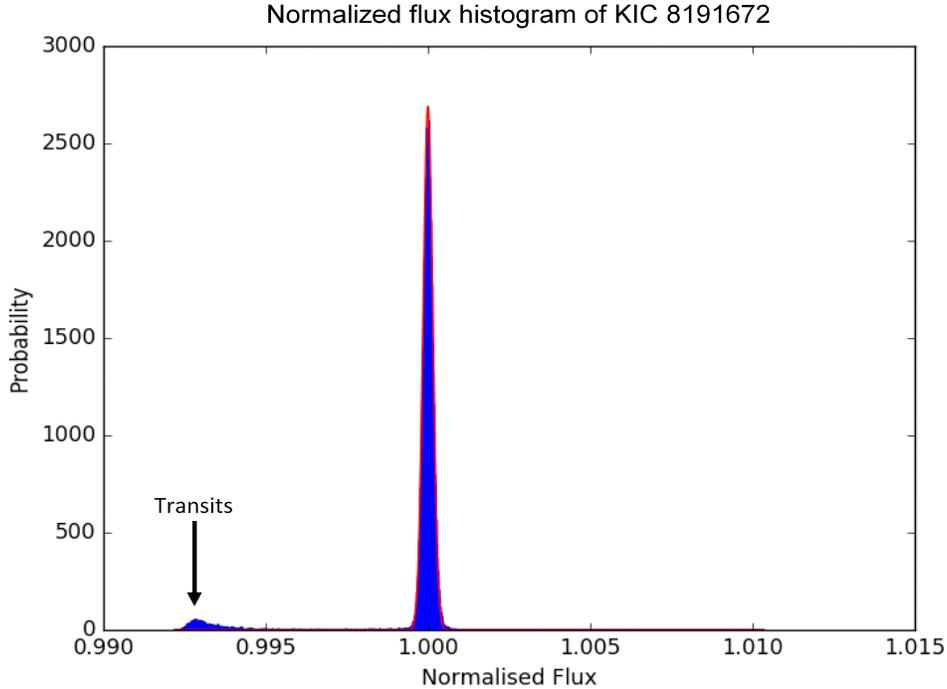
Figure 2.2: Normalized flux histogram and Gaussian distribution of
Kepler-5 b (KIC 8191672) after $3 - \sigma$ clipping.

rical, it would indicate the 1σ flux error measurement was unreliable. We
assume that stars with no companions/transiting planets are consistent with
white noise (a Gaussian/normal distribution) which is symmetric and yields
a skewness of 0. The exoplanet transits should lie outside the 3σ clipping (see
Figure 2.2) and hence not be included in the estimate for the 1σ flux error,
leading to a symmetric distribution consistent with white noise. Kepler-5 b
(Figures 2.2, 2.3 and 2.4) is an exoplanet that is part of our algorithm testing
and calibration sample (refer to Test Sample subsection).

We modelled exoplanet transit light curves using PyTransit (Parviainen [2015]). The model parameters are the planet-star radius ratio, k, the transit center, t_0, the orbital period, p, the scaled semi-major axis (in units of solar radii), a, the orbital inclination, i, the orbital eccentricity, e, the argument of periastron, w, and the quadratic limb darkening coefficients, u.

Limb darkening is a phenomenon that results in the stellar disk being brighter in the middle and fainter at the edge i.e. the limb (Dotson et al. [2010]). We set the limb darkening coefficients to 0 (no limb darkening) because there is evidence to show that concurrent limb darkening tables do not yield consistent results and these inconsistencies will lead to the incorrect determination of the planet-star radius ratio (Csizmadia et al. [2013]).

We assumed the exoplanets had circular ($e = 0$ and $w = 0$) and edge-on ($i = 90°$) orbits as observed from Earth. We implemented the Mandel and Agol [2002] transit model but assumed no limb darkening instead of quadratic limb darkening. We fitted for a single planet in each light curve (multi-planetary systems are discussed in Chapter 4). The output of the transit model is normalized flux which is plotted over a chosen time interval. We fitted for four model parameters, namely: k, the planet-star radius ratio, t_0, the transit center, p, the orbital period, and the scaled semi-major axis, a.

We performed a non-linear least-squares fit of the model to the data using a chi-square minimization approach to get the best fit parameters of the model. The chi-square statistic is defined for a non-linear least-squares

model-data fitting problem as,

$$\chi^2 = \sum_i^N \left(\frac{[y_i^{data} - y_i^{model}(n)]^2}{\epsilon_i^2} \right) \tag{2.1}$$

where y_i^{data} are the data points, y_i^{model} is the model calculation, n are the set of model parameters to be fitted for and ϵ_i is the uncertainty in the data (Newville et al. [2014]).

We used the Powell method for the chi-square minimization as it does not require the chi-square function to be differentiable, and no derivatives are taken. The Powell method finds the minimum of a function of n parameters $x_1, x_2, ..., x_n$ such that the value of the function at these parameters $f(x_1, x_2, ..., x_n)$ is a minimum. In our case, we fitted for $n = 4$ parameters and our function to be minimized was equation (2.1) for the chi-squared statistic (Powell [1964]). The LMFIT (Non-Linear Least-Squares Minimization and Curve-Fitting for Python) python package was used to minimize equation (2.1) (Newville et al. [2014]).

We found that the chi-square function had several local minima corresponding to different period values (see Figure 2.3). Unlike the Levenberg-Marquardt method (Levenberg [1944], Marquardt [1963]) which is favoured for non-linear least-squares model-data fitting problems and requires differentiable functions and derivatives, the Powell method does not require derivatives (gradients) to be calculated to find the function minimum. Thus, we found it less likely for the minimization algorithm to get trapped in regions
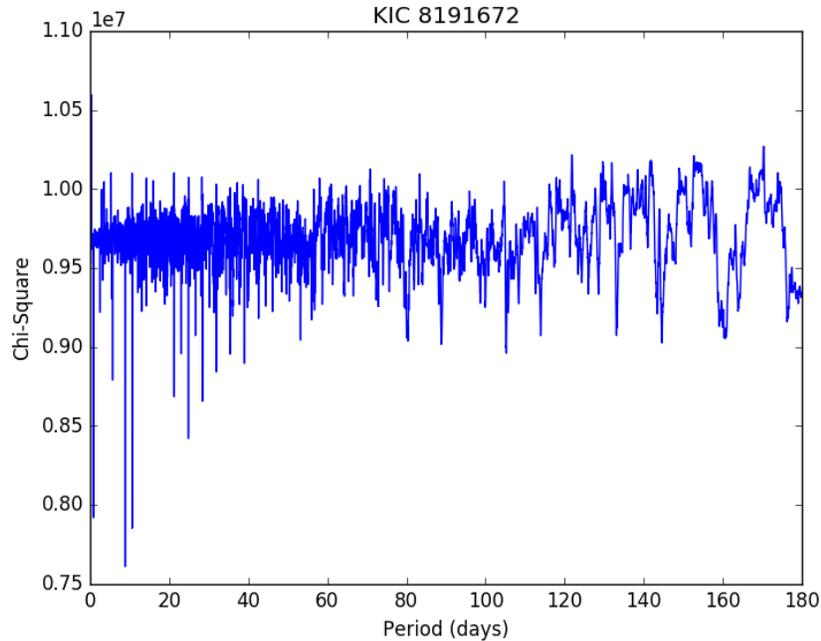
Figure 2.3: Plot of chi-square as a function of period for Kepler-5 b (KIC 8191672) to illustrate regions of local minima.

of local minima and yield the incorrect best fit parameters when using the Powell method. Despite our Powell method precautionary measure, the minimization algorithm did get trapped in regions of local minima which required us to use a semi-brute force approach.

We placed bounds on the parameters $(k, \text{t\_}0, p, a)$ to ensure logical and physically meaningful values for the best fit parameter values:

1. To find the transit centre, t_0, which corresponds to a position on the time axis, we looped through the 10 minimum flux values in the light curve to ensure the transit centre value was not disturbed from the true

value by any random outliers. This is a semi-brute force approach.

2. To calculate the orbital period, we looped through period values from 0 to 180 days and set minimum and maximum bounds for the period around a central period value, based on the period bin edges in the range 0-180 days. This was a precaution to ensure the minimization algorithm did not get trapped in regions of local minima. We did not anticipate finding exoplanets with orbital periods longer than 180 days with Kepler long cadence light curves due to the 3 month observation window and the fact that we were using two quarters of data (equivalent to 180 days of observation).

3. $0 \leq k \leq 1$.

   $k = 0$ corresponds to fitting a flat-line model to the data i.e. there is no change in stellar flux and hence we assumed there were no transiting planets/companions. For $k = 0$, all the other parameter values become meaningless.

4. $1R_{\odot} \leq a$

   We assumed it unlikely to have planet-star separations less than $1R_{\odot}$ and let the maximum value for the semi-major axis be unbound.

Using the bounds and conditions listed above, the chi-square values were recorded for each iteration and the best fit parameters, which were those parameter values that corresponded to the minimum chi-square statistic value,

were recorded for each star.

The 1σ uncertainty in the model best fit parameters (1σ confidence limits) were calculated using the $\Delta\chi^2$ method with 4 degrees of freedom (4 parameters). If the parameter values are perturbed away from their best fit values, χ^2 increases. The region within which χ^2 increases by no more than a set amount $\Delta\chi^2$ defines some n-dimensional confidence region around the best fit parameters, where $n = 4$ in this case (Press et al. [1989]).

The region of confidence (with significance level α) is defined by

$$\chi^2_\alpha = \chi^2_{\text{min}} + \Delta\chi^2(n, \alpha) \ . \tag{2.2}$$

So, in order to estimate the 1σ ($\alpha = 0.68$) uncertainty in $n = 4$ best fit parameter values

$$\chi^2_{0.68} = \chi^2_{\text{min}} + 4.72 \ . \tag{2.3}$$

Finally, we generated phased light curve plots of normalized flux vs period, centred around the exoplanet transit. We plotted the model, with best fit parameters calculated from the chi-square minimization, and the data on the same set of axes (see Figure 2.4).
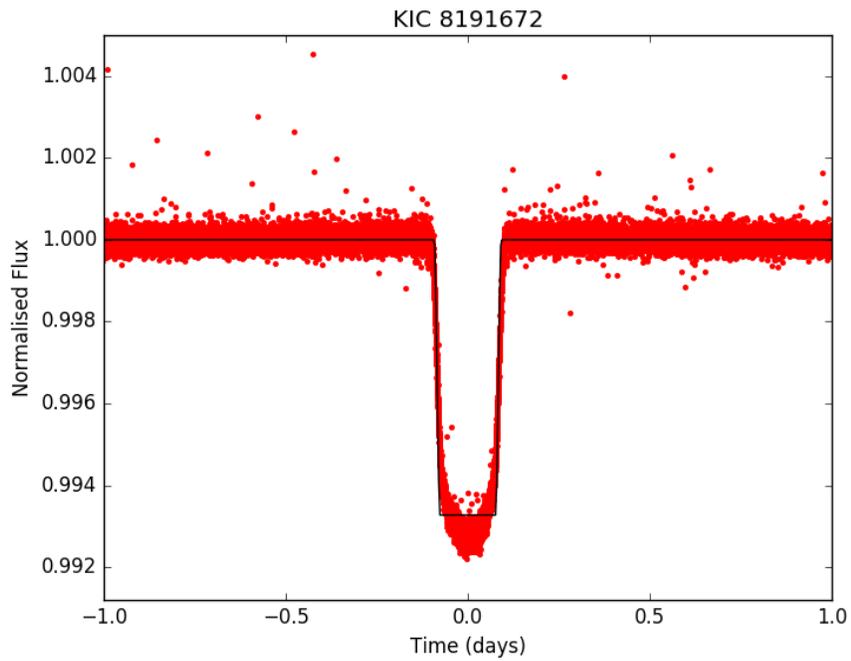
Figure 2.4: Phased transit light curve and best fit parameter model (solid black line) of Kepler-5 b (KIC 8191672). The algorithm measured the period to be $p = 3.54847 \pm 0.00001$ days, with $k = 0.08201 \pm 0.00004$. The literature period for Kepler-5 b is 3.5484657 ± 0.0000007 days.

2.2.1 Test Sample

To test and calibrate our exoplanet detection algorithm, we used a test sample of Kepler data consisting of 10 randomly selected stars with confirmed exoplanets. The exoplanets in the test sample were all from single planetary systems. Tables 2.1-2.3 compare the calculated parameter values and the literature parameter values for the exoplanets in the test sample. Table A3 in the Appendix lists the test sample exoplanets with their associated star Kepler Input Catalog identifications (KIC IDs) and planet-star radius ratios.

Three of the exoplanets in the test sample, Kepler-439 b ($p = 178.1396^{+0.0016}_{-0.0018}$ days), Kepler-441 b ($p = 207.2482^{+0.0022}_{-0.0020}$ days) and Kepler-443 b ($p = 177.6693^{+0.0031}_{-0.0030}$ days) had planet-star radius ratios of 0 which made the other parameter values meaningless hence they are not included in Tables 2.1-2.3. Kepler-439 b, Kepler-441 b and Kepler-443 b are all long-period exoplanets and the algorithm did not identify any distinct transits and flat models were fitted. Kepler-441 b's period of 207.2482 days is out of the period range of the algorithm, hence it went undetected.

We found the calculated orbital period values were within $\pm 3\sigma$ of the literature orbital period values for Kepler-5 b, Kepler-12 b, Kepler-434 b, Kepler-410A b, Kepler-77 b and Kepler-8 b.

Kepler-78 b was an extremely short-period exoplanet ($p = \pm 9$ hours) and a small planet-star radius ratio ($k \approx 0.002$), hence our algorithm did not calculate the parameter values correctly for this extreme case.

Kepler Name	Calculated Orbital Period (days)	Literature Orbital Period (days)
Kepler-5 b	3.54847 ± 0.00001	3.5484657 ± 0.0000007
Kepler-78 b	6.09311 ± 0.00035	0.3550 ± 0.0004
Kepler-12 b	4.437962 ± 0.000001	4.4379629 ± 0.0000006
Kepler-434 b	12.87597 ± 0.00005	12.8747099 ± 0.0000050
Kepler-410A b	17.8304 ± 0.0039	17.833648 ± 0.000054
Kepler-77 b	3.57874 ± 0.00001	$3.57878087 \pm 0.00000023$
Kepler-8 b	3.52250 ± 0.00001	3.5224991 ± 0.0000007

Table 2.1: Calculated vs. literature orbital periods of exoplanets from test sample.

Kepler Name	Calculated Planet Radius (R_{Jup})	Literature Planet Radius (R_{Jup})
Kepler-5 b	1.4314 ± 0.0007	$1.426^{+0.036}_{-0.051}$
Kepler-78 b	0.0210 ± 0.0017	$0.105^{+0.014}_{-0.008}$
Kepler-12 b	1.77692 ± 0.00058	$1.754^{+0.031}_{-0.036}$
Kepler-434 b	0.7607 ± 0.0095	$1.13^{+0.26}_{-0.18}$
Kepler-410A b	0.0442 ± 0.0014	0.253 ± 0.005
Kepler-77 b	0.9766 ± 0.0013	0.960 ± 0.016
Kepler-8 b	1.31629 ± 0.00087	$1.416^{+0.053}_{-0.062}$

Table 2.2: Calculated vs. literature planetary radius of exoplanets from test sample.

Kepler Name	Calculated Semi-Major Axis (AU)	Literature Semi-Major Axis (AU)
Kepler-5 b	0.031260 ± 0.000015	$0.0538^{+0.0015}_{-0.0021}$
Kepler-78 b	0.00516 ± 0.00012	N/A
Kepler-12 b	0.039606 ± 0.000012	$0.0553^{+0.0010}_{-0.0012}$
Kepler-434 b	0.3233 ± 0.0031	0.1143 ± 0.0030
Kepler-410A b	0.0071 ± 0.00015	0.1226 ± 0.0047
Kepler-77 b	0.051159 ± 0.000076	0.04501 ± 0.00063
Kepler-8 b	0.04733 ± 0.00003	$0.0474^{+0.0018}_{-0.0021}$

Table 2.3: Calculated vs. literature semi-major axis of exoplanets from test sample.

The calculated planetary radius values were within $\pm 3\sigma$ of the literature planetary radius values for Kepler-5 b, Kepler-12 b, Kepler-77 b and Kepler-8 b. These planets are all short-period hot Jupiters with large planet-star radius ratios ($k \approx 0.01$), leading our algorithm to fit correctly for the transit depth and calculate the planetary radii correctly.

Kepler-410Ab is a $2.8R_{\oplus}$ exoplanet orbiting the binary star system Kepler-410A. Though its orbital period was calculated correctly, its planetary radius and semi-major axis were incorrect and this could be due to interference from the binary star system in the light curve data.

All of the test sample planets (with the exception of Kepler-439 b, Kepler-441 b and Kepler-443 b) had $k > 0.001$.

There were some inconsistencies in determining the semi-major axis values, a, and this could be due the fact that the mass of the star (see equation (1.4)) is not available to the PyTransit package used to model the light curves and calculate the semi-major axis values.

2.2.2 Blind Sample

Once our exoplanet detection algorithm was tested and calibrated, we applied it to a blind sample of Kepler mission data consisting of approximately 4500 stars. We used the University of KwaZulu-Natal's high performance computing facility, Hippo, to run the data through the algorithm.

We determined the properties of the exoplanets found in the blind sample

based on the best fit parameters calculated using our exoplanet detection algorithm. The size of the planet is determined from the planet-star radius ratio and the scaled semi-major axis parameter is used to determine if the planet falls within its star's habitable zone. The habitable zone boundaries are calculated by the NASA Exoplanet Archive Predicted Observables for Exoplanets Service[2], based on calculations from Kasting et al. [1993]. The planet size (radius) and the semi-major axis are used to determine if the exoplanet is potentially habitable.

Our exclusion criteria for light curves we deemed did not contain exoplanet transits were $k < 0.001$ and $p < 0.5$ days. The Earth orbiting the Sun (a G-class star) has a planet-star radius ratio of $k = 0.01$. Since our blind sample contained Sun-like stars and smaller, cooler stars i.e. G, K and M class stars, and we were searching for habitable, Earth-like planets, we did not consider planet-star radius ratios less than 0.001. Once our exclusion criteria were met, these light curves were excluded from uncertainty estimations and phased light curve plot generation.

In order to make computing more feasible for stars with $k > 0.001$ and $p > 0.5$ days, parameter error values were recorded as 0.0 if the error values calculated using the $\Delta\chi^2$ method did not converge within 200 seconds. There were 3 instances of this for calculated orbital period (Table 3.1), 5 instances for calculated planet radius (Table 3.2) and 11 instances for calculated semi-major axis (Table 3.3).

---

[2]https://exoplanetarchive.ipac.caltech.edu/cgi-bin/CalcQty/nph-calcqty

Chapter 3

Results

In this Chapter, we define the blind sample of Kepler mission data used in this dissertation and present our results for the exoplanets our algorithm detected in the blind sample. We compare the calculated parameter values to the literature parameter values for the detected exoplanets.

The criteria for determining the blind sample were:

1. $T_{star} < 6000$ K

2. $R_{star} < 1R_{\odot}$

3. $13.5 <$ Kepler Magnitude < 14

We determined the criteria by selecting stars we thought we would have a better chance of detecting exoplanets around. We selected Sun-like stars

and smaller, lower mass stars ($T_{star} < 6000$ K, $R_{star} < 1R_\odot$) to ensure large planet-star radius ratios and hence more probable detections. The Kepler magnitude criterion was arbitrary in order to keep the number of objects at a reasonable level for the purposes of this dissertation.

We used Kepler mission data from quarters 5 and 6 in 2010 i.e. Q5 and Q6. This corresponds to approximately 180 observation days between 20 March 2010 and 23 September 2010.

The blind sample consisting of 4503 stars contained 70 known exoplanets. Our exoplanet detection algorithm found 50 of these 70 planets and the results are presented in Tables 3.1-3.3. Table 3.1 compares the calculated orbital period and literature orbital period values of the 50 exoplanets found. Table 3.2 compares the calculated and literature planetary radii values. Table 3.3 compares the calculated and literature semi-major axis values[1]. The literature values are obtained from the NASA Exoplanet Archive of confirmed exoplanets[2].

21 of the 50 detected exoplanets (42%) had calculated orbital period values within $\pm 3\sigma$ of the literature orbital period values. 13 of the 50 detected exoplanets (26%) had calculated planetary radius values within $\pm 3\sigma$ of the literature planetary radius values. Only 19 of the 50 detected exoplanets have literature semi-major axis values, hence calculated and literature semi-

---

[1] A large number of exoplanets do not have literature semi-major axis values as follow-up observations are still pending to confirm these values.

[2] https://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/nph-tblView?app=ExoTbls&config=planets

major axis comparisons cannot be made for 31 of the 50 detected exoplan-
ets. 3 detected exoplanets of the 19 with literature semi-major axis values
(15.8%) had calculated semi-major axis values within $\pm 3\sigma$ of the literature
semi-major axis values. Explanations and shortcomings for these results are
discussed in Chapter 4.

Figure 3.1 is a plot of calculated and literature periods of the detected
exoplanets as a function of radius, showing the sensitivity of the algorithm,
and indeed the transit method, toward detecting short-period exoplanets
larger than the Earth as the majority of planets found in the sample were
super Earths, Neptune-sized exoplanets and a few hot Jupiters. This is
discussed further in Chapter 4.

A full list of the exoplanets found by our algorithm in this sample, with
their associated star KIC IDs and planet-star radius ratios, can be found in
Appendix, Table A1.
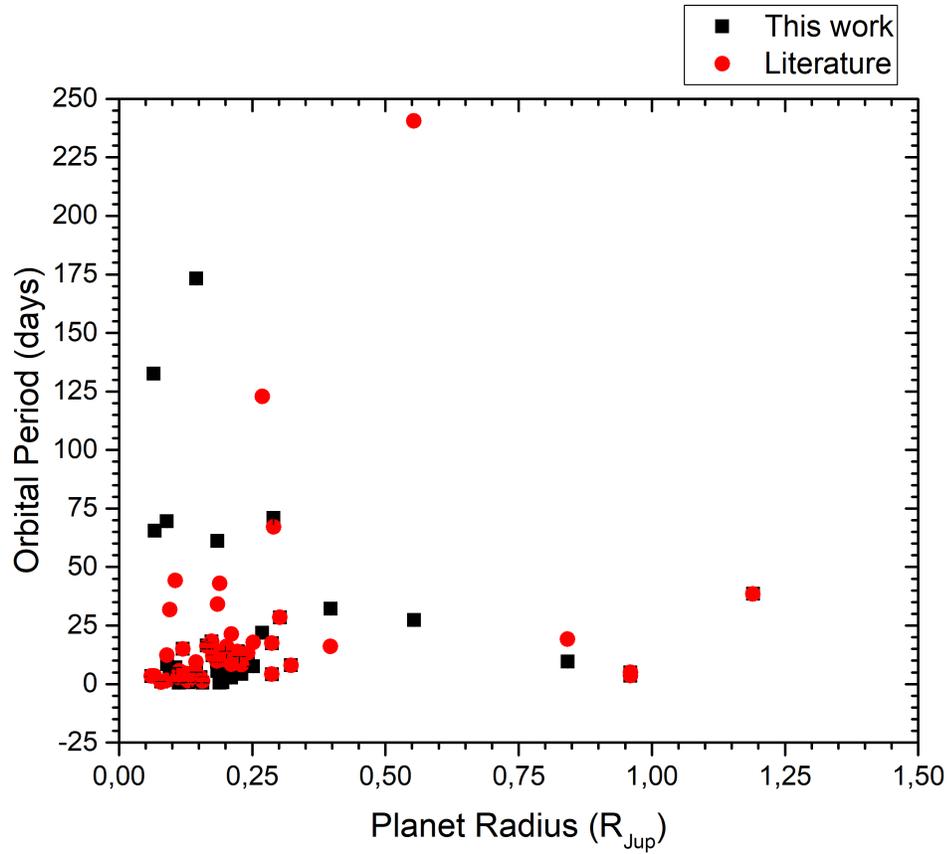
Figure 3.1: Plot of calculated period and literature period as a function of exoplanet radius.

Kepler Name	Calculated Orbital Period (days)	Literature Orbital Period (days)
Kepler-261 b	0.69228 ± 0.00004	10.381227 ± 0.000016
Kepler-479 b	2.50024 ± 0.00006	$12.49341307 \pm 0.0000136$
Kepler-9 b	9.63983 ± 0.00006	19.24 ± 0.0
Kepler-198 b	7.5456 ± 0.0015	17.790037 ± 0.000035
Kepler-543b	13.8996 ± 0.0001	$13.89961966 \pm 0.00001102$
Kepler-472 b	4.17551 ± 0.00004	$4.176255514 \pm 0.000001844$
Kepler-319 b	173.21 ± 0.15	4.362705 ± 0.000009
Kepler-199 c	70.8767 ± 0.0018	67.093408 ± 0.000190
Kepler-964 b	5.1990 ± 0.0030	$13.5225106 \pm 0.00002258$
Kepler-636 b	32.15853 ± 0.00014	16.08066 ± 0.000005603
Kepler-755 b	0.63455 ± 0.00001	1.26909 ± 0.000001005
Kepler-202 c	16.52 ± 1.53	16.282493 ± 0.000038
Kepler-1499 b	7.0633 ± 0.0070	44.2008 ± 0.0006553
Kepler-775 b	0.50034 ± 0.0004	$0.974868926 \pm 0.0000008$
Kepler-205 b	1.37739 ± 0.00002	2.755640 ± 0.000002
Kepler-477 b	5.55806 ± 0.00003	$11.11990653 \pm 0.0000116$
Kepler-209 b	13.90363 ± 0.00028	16.087845 ± 0.000041
Kepler-831 b	0.56904 ± 0.00012	$5.62153941 \pm 0.00001272$
Kepler-210 c	7.97276 ± 0.00003	7.972513 ± 0.000003
Kepler-112 b	8.40913 ± 0.00011	8.408878 ± 0.000010
Kepler-322 b	69.4914 ± 0.0020	1.653888 ± 0.000006
Kepler-77 b	3.57874 ± 0.00001	$3.57878087 \pm 0.00000023$
Kepler-69 b	13.72385 ± 0.00042	$13.722341^{+0.000035}_{-0.000036}$
Kepler-648 b	17.41993 ± 0.00009	$17.4211749 \pm 0.00002066$
Kepler-215 b	5.13630 ± 0.00042	9.360672 ± 0.000040
Kepler-62 d	18.1685 ± 0.0002	18.16406 ± 0.00002
Kepler-1173 b	1.01427 ± 0.00039	$0.7698536 \pm 0.000001205$
Kepler-1178 b	6.5045 ± 0.0018	31.80634 ± 0.0003857
Kepler-1114 b	15.03767 ± 0.0	$14.97435694 \pm 0.00005221$
Kepler-958 b	9.529 ± 0.016	9.7678805 ± 0.0000111
Kepler-326 c	4.6792 ± 0.0034	4.580358 ± 0.000010
Kepler-461 b	4.33973 ± 0.00054	$8.313783059 \pm 0.000007661$
Kepler-478 b	13.22087 ± 0.00006	13.2217576 ± 0.0000139
Kepler-651 b	2.67301 ± 0.00005	$21.38521506 \pm 0.0000388$

Table 3.1: Calculated vs. literature orbital periods of exoplanets detected using our algorithm.

Kepler Name	Calculated Orbital Period (days)	Literature Orbital Period (days)
Kepler-962 b	12.189 ± 0.014	$12.05707239 \pm 0.00001586$
Kepler-453 b	27.32248 ± 0.00003	240.503 ± 0.053
Kepler-984 b	0.50022 ± 0.00035	43.0342272 ± 0.0001479
Kepler-221 b	2.84526 ± 0.00002	2.795906 ± 0.000004
Kepler-1558 b	13.3889 ± 0.0064	$3.50470358 \pm 0.00001612$
Kepler-1019 b	0.70561 ± 0.00001	$1.411229847 \pm 0.000000746$
Kepler-1371 b	132.57541 ± 0.0	3.4462039 ± 0.00002017
Kepler-969 b	61.1329 ± 0.0017	$34.1731714 \pm 0.00007358$
Kepler-114 d	11.77504 ± 0.00007	11.776 ± 0.0
Kepler-532 b	12.9410 ± 0.0011	$12.92491623 \pm 0.0000123$
Kepler-1482 b	8.564 ± 0.029	$12.25383217 \pm 0.00008407$
Kepler-533 b	28.50697 ± 0.00033	$28.51120525 \pm 0.00002321$
Kepler-15 b	4.94277 ± 0.00001	4.942782 ± 0.0000013
Kepler-468 b	38.48180 ± 0.00007	$38.478757067 \pm 0.000004851$
Kepler-1036 b	21.83 ± 0.11	122.8808058 ± 0.000709
Kepler-1563 b	65.5 ± 0.0	$3.43276598 \pm 0.00002976$

Table 3.1 continued...

Kepler Name	Calculated Planet Radius (R_{Jup})	Literature Planet Radius (R_{Jup})
Kepler-261 b	0.0562 ± 0.0064	0.194 ± 0.012
Kepler-479 b	0.0728 ± 0.0067	$0.189^{+0.008}_{-0.006}$
Kepler-9 b	0.5782 ± 0.0041	0.842 ± 0.069
Kepler-198 b	0.0712 ± 0.0064	0.252 ± 0.037
Kepler-543b	0.1892 ± 0.0041	$0.223^{+0.011}_{-0.023}$
Kepler-472 b	0.2569 ± 0.0043	$0.287^{+0.012}_{-0.025}$
Kepler-319 b	0.267 ± 0.013	0.145 ± 0.021
Kepler-199 c	0.292 ± 0.013	0.290 ± 0.052
Kepler-964 b	0.0743 ± 0.0061	$0.189^{+0.008}_{-0.006}$
Kepler-636 b	0.4026 ± 0.0053	0.397 ± 0.012
Kepler-755 b	0.0707 ± 0.0052	$0.157^{+0.005}_{-0.028}$
Kepler-202 c	0.018 ± 0.0	0.165 ± 0.009
Kepler-1499 b	0.0511 ± 0.0062	$0.106^{+0.022}_{-0.012}$
Kepler-775 b	0.0190 ± 0.0070	0.106 ± 0.004
Kepler-205 b	0.0790 ± 0.0037	0.135 ± 0.012
Kepler-477 b	0.1299 ± 0.0055	$0.185^{+0.016}_{-0.011}$
Kepler-209 b	0.1394 ± 0.0068	0.202 ± 0.030
Kepler-831 b	0.0263 ± 0.0087	$0.113^{+0.009}_{-0.006}$
Kepler-210 c	0.3157 ± 0.0027	0.323 ± 0.017
Kepler-112 b	0.2187 ± 0.0044	0.211 ± 0.051
Kepler-322 b	0.192 ± 0.017	0.090 ± 0.010
Kepler-77 b	0.9766 ± 0.0013	0.960 ± 0.016
Kepler-69 b	0.1903 ± 0.0063	$0.200^{+0.039}_{-0.026}$
Kepler-648 b	0.2527 ± 0.0068	$0.287^{+0.011}_{-0.021}$
Kepler-215 b	0.067 ± 0.010	0.145 ± 0.036
Kepler-62 d	0.1738 ± 0.0050	0.174 ± 0.006
Kepler-1173 b	0.0261 ± 0.0068	$0.0794^{+0.0205}_{-0.0062}$
Kepler-1178 b	0.0409 ± 0.0079	$0.0955^{+0.0054}_{-0.0071}$
Kepler-1114 b	0.0081 ± 0.0	$0.120^{+0.007}_{-0.006}$
Kepler-958 b	0.0379 ± 0.0064	$0.184^{+0.009}_{-0.007}$
Kepler-326 c	0.0542 ± 0.0048	0.125 ± 0.009
Kepler-461 b	0.0812 ± 0.0049	$0.230^{+0.006}_{-0.005}$
Kepler-478 b	0.2142 ± 0.0072	$0.242^{+0.008}_{-0.029}$
Kepler-651 b	0.0691 ± 0.0086	$0.211^{+0.020}_{-0.012}$

Table 3.2: Calculated vs. literature planetary radii of exoplanets detected using our algorithm.

Kepler Name	Calculated Planet Radius (R_{Jup})	Literature Planet Radius (R_{Jup})
Kepler-962 b	0.2118 ± 0.0049	0.176 ± 0.007
Kepler-453 b	2.1860 ± 0.0018	0.5535 ± 0.0035
Kepler-984 b	0.0225 ± 0.0091	$0.189^{+0.012}_{-0.010}$
Kepler-221 b	0.1815 ± 0.0014	0.153 ± 0.015
Kepler-1558 b	0.0524 ± 0.0043	$0.061^{+0.005}_{-0.004}$
Kepler-1019 b	0.0777 ± 0.0036	$0.130^{+0.004}_{-0.014}$
Kepler-1371 b	0.0084 ± 0.0	$0.065^{+0.007}_{-0.004}$
Kepler-969 b	0.160 ± 0.012	0.185 ± 0.004
Kepler-114 d	0.2264 ± 0.0040	0.226 ± 0.025
Kepler-532 b	0.0945 ± 0.0071	$0.229^{+0.019}_{-0.012}$
Kepler-1482 b	0.0254 ± 0.0083	$0.0901^{+0.0089}_{-0.0098}$
Kepler-533 b	0.2627 ± 0.0045	$0.302^{+0.007}_{-0.029}$
Kepler-15 b	0.9647 ± 0.0013	$0.96^{+0.06}_{-0.07}$
Kepler-468 b	1.2453 ± 0.0022	$1.190^{+0.029}_{-0.026}$
Kepler-1036 b	0.040 ± 0.0	$0.269^{+0.012}_{-0.047}$
Kepler-1563 b	0.013 ± 0.0	$0.067^{+0.011}_{-0.007}$

Table 3.2 continued...

Kepler Name	Calculated Semi-Major Axis (AU)	Literature Semi-Major Axis (AU)
Kepler-261 b	0.00793 ± 0.00027	0.088
Kepler-479 b	0.02651 ± 0.00036	N/A
Kepler-9 b	0.0961 ± 0.0004	0.140 ± 0.001
Kepler-198 b	0.00891 ± 0.00036	0.131
Kepler-543b	0.1568 ± 0.0016	N/A
Kepler-472 b	0.04572 ± 0.00022	N/A
Kepler-319 b	0.52 ± 0.0	0.051
Kepler-199 c	0.48 ± 0.0	0.316
Kepler-964 b	0.00675 ± 0.00024	N/A
Kepler-636 b	0.5146 ± 0.0037	N/A
Kepler-755 b	0.01224 ± 0.00017	N/A
Kepler-202 c	0.0050 ± 0.0	0.113
Kepler-1499 b	0.00697 ± 0.00022	N/A
Kepler-775 b	0.0056 ± 0.0011	N/A
Kepler-205 b	0.03070 ± 0.00055	0.032
Kepler-477 b	0.06721 ± 0.00029	N/A
Kepler-209 b	0.0647 ± 0.0012	0.122
Kepler-831 b	0.0055 ± 0.0020	N/A
Kepler-210 c	0.1111 ± 0.0010	0.070
Kepler-112 b	0.0945 ± 0.0016	0.076
Kepler-322 b	0.62 ± 0.0	0.027
Kepler-77 b	0.051159 ± 0.000076	0.04501 ± 0.00063
Kepler-69 b	0.1038 ± 0.0015	$0.094^{+0.023}_{-0.016}$
Kepler-648 b	0.12768 ± 0.00095	N/A
Kepler-215 b	0.02770 ± 0.00065	0.084
Kepler-62 d	0.2167 ± 0.0030	0.120 ± 0.001
Kepler-1173 b	0.00477 ± 0.00013	N/A
Kepler-1178 b	0.01017 ± 0.00093	N/A
Kepler-1114 b	0.0047 ± 0.0	N/A
Kepler-958 b	0.00588 ± 0.00012	N/A
Kepler-326 c	0.00633 ± 0.00018	0.051
Kepler-461 b	0.006951 ± 0.000088	N/A
Kepler-478 b	0.1623 ± 0.0011	N/A
Kepler-651 b	0.02746 ± 0.00046	N/A

Table 3.3: Calculated vs. literature semi-major axis of exoplanets detected using our algorithm.

Kepler Name	Calculated Semi-Major Axis (AU)	Literature Semi-Major Axis (AU)
Kepler-962 b	0.0742 ± 0.0016	N/A
Kepler-453 b	0.23253 ± 0.00026	0.7903 ± 0.0028
Kepler-984 b	0.0054 ± 0.0020	N/A
Kepler-221 b	0.0482000 ± 0.0000056	0.037
Kepler-1558 b	0.00928 ± 0.00057	N/A
Kepler-1019 b	0.01764 ± 0.00046	N/A
Kepler-1371 b	0.0075 ± 0.0	N/A
Kepler-969 b	0.29 ± 0.0	N/A
Kepler-114 d	0.1566 ± 0.0011	N/A
Kepler-532 b	0.050 ± 0.0	N/A
Kepler-1482 b	0.0055 ± 0.0	N/A
Kepler-533 b	0.2168 ± 0.0022	N/A
Kepler-15 b	0.06588 ± 0.00014	$0.05714^{+0.00086}_{-0.00093}$
Kepler-468 b	0.33408 ± 0.00065	N/A
Kepler-1036 b	0.0088 ± 0.0	N/A
Kepler-1563 b	0.0047 ± 0.0	N/A

Table 3.3 continued...

Chapter 4

Discussion

Our blind sample consisting of 4503 stars contained 70 known exoplanets. In this Chapter, we will discuss the exoplanets detected by our algorithm in the blind sample, as well as the exoplanets that went undetected. An exoplanet is considered detected if $k > 0.001$ and $p > 0.5$ days, and undetected if $k < 0.001$. We will discuss:

1. Detected exoplanets with correctly calculated parameter values as compared with the literature,

2. Detected exoplanets with incorrectly calculated parameter values as compared with the literature and,

3. Undetected exoplanets in the blind sample.

We will explore various factors associated with each scenario and provide possible suggestions and improvements.

4.1　Detected Exoplanets

Our exoplanet detection algorithm found 50 of the 70 exoplanets in the sample and 20 of those 50 planets were from multi-planetary systems.

4.1.1　Accurate Parameter Estimations

Correctly Calculated Orbital Period and Planetary Radius

13 of the 50 planets found had calculated orbital period and planetary radius values within $\pm 3\sigma$ of the literature values listed in the NASA Exoplanet Archive[1]. These 13 planets either had radii $> 2.5 R_{\oplus}$ and/or were orbiting smaller stars (K and M class) resulting in planet-star radius ratios larger than 0.01. For $k > 0.01$, the transit centre, and thus the transit depth, can be fitted more accurately. We found that all planets in the literature with $k > 0.01$ from single planetary systems were successfully recovered by our algorithm, and their orbital period and planetary radius values as measured by our pipeline are in agreement with those reported previously in the literature (see Tables 3.1, 3.2 and A1), provided that the light curves do not exhibit harmonics (see section 4.1.2 - Period Harmonics) and have been correctly cleaned of stellar/systematic/random variations (see section 4.1.2 - Median Filtering).

Among these 13 planets, 3 were hot Jupiters (planets with short orbital

---

[1]https://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/nph-tblView?app=ExoTbls&config=planets

periods and roughly the size of Jupiter) orbiting Sun-like stars $(0.8R_\odot - 1R_\odot)$ and were from single planetary systems.

Studying the light curves of these 3 hot Jupiters (Kepler-15 b, Kepler-77 b and Kepler-468 b), we saw a transit centered directly on the transit center, and a model fitted almost perfectly to the transit curve. The transit depth (dip) is fitted almost entirely, leading to the correct determination of the planet-star radius ratio and hence the planet radius (see Figure 4.1 of Kepler-77 b). There were some inconsistencies in determining the semi-major axis value, a, and this could be due the fact that the mass of the star (see equation (1.4)) is not available to the PyTransit package used to model the light curves and calculate the semi-major axis value.

4 of the 13 planets with calculated orbital period and planetary radius values within $\pm 3\sigma$ of the literature values were short period planets ($p <$ 20 days) with radii $> 2.5R_\oplus$ (or $R_{planet} > 0.2R_{Jup}$), somewhat larger than super Earths, in single planetary systems orbiting stars with radii between $0.67 - 0.9R_\odot$.

The light curves of these 4 exoplanets (Kepler-472 b, Kepler-478 b, Kepler-543 b and Kepler-648 b), have transits centered directly on the transit center, and a model fitted almost perfectly to the transit curve. The transit depth, though not as large for the hot Jupiter in Figure 4.1, is fitted almost entirely, leading to the correct determination of the planet-star radius ratio and hence the planet radius (see Figure 4.2 of Kepler-472 b).
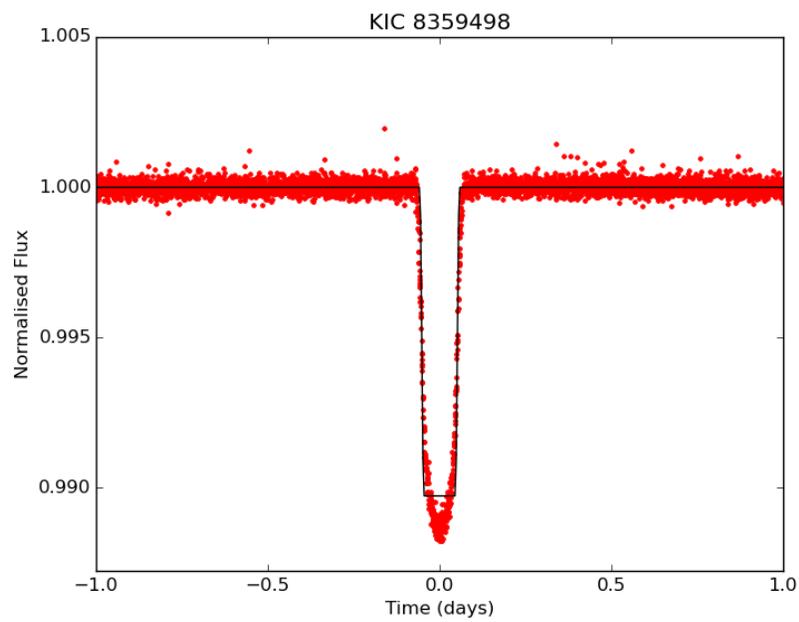
Figure 4.1: Phased transit light curve and best fit parameter model of hot Jupiter Kepler-77 b (KIC 8359498) showing well-fitted transit depth (solid black line).
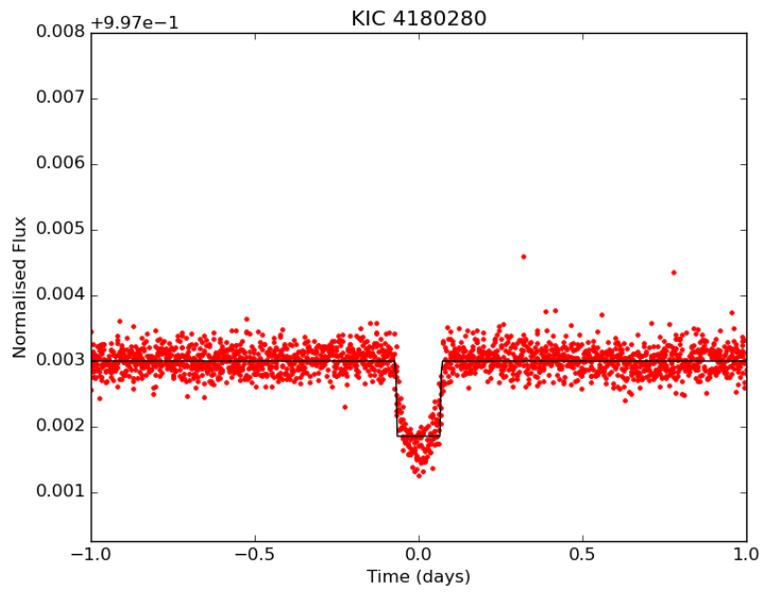
Figure 4.2: Phased transit light curve and best fit parameter model of Neptune-sized Kepler-472 b (KIC 4180280) showing well-fitted transit depth.

The remaining 6 planets of the 13 exoplanets with calculated orbital period and planetary radius values within $\pm 3\sigma$ of the literature values (Kepler-62 d, Kepler-69 b, Kepler-112 b, Kepler-114 d, Kepler-199 b and Kepler-210 c) were from multi-planetary systems.

These 6 planets were the largest planets in their planetary systems and had orbital periods of less than 20 days. The Kepler-112 system has 2 exoplanets with the same radius so the algorithm fitted for the shorter period exoplanet (Kepler-112 b, $p = 8.4$ days - see Figure 4.3) instead of the longer period planet ($p = 28.6$ days).

The calculated orbital period and planetary radius values were within $\pm 3\sigma$ of the literature values because the planets in these multi-planetary systems had distinct period values (e.g. the 2 planets in the Kepler-210 system had periods of 2.45 and 7.97 days), hence the model fitting was not disturbed by overlapping transits.

Our exoplanet search algorithm first fits for the largest transit depth and/or (subsequently) for the shortest orbital period.

Kepler-62 is a K-class star, host to 5 exoplanets - 4 super Earths, including Kepler-62 d, and 1 planet with radius=$0.5R_\oplus$. Kepler-62 e has a period of 122.4 days which may not have been entirely captured in the 180 days of observations used, and Kepler-62 f has a period of 267 days which is outside the period search range of our algorithm. Kepler-62 e and Kepler-62 f are potentially habitable super Earths. Kepler-62 e is part of the optimistic sample of potentially habitable exoplanets (Figure 1.14) and Kepler-62 f is
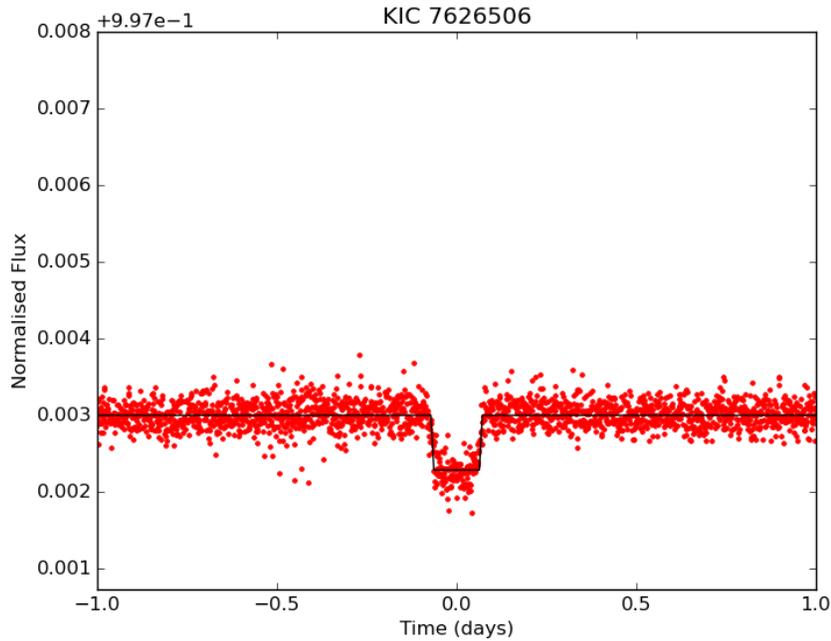
Figure 4.3: Phased transit light curve and best fit parameter model of Kepler-112 b (KIC 7626506) from multi-planetary system with $p = 8.4$ days, showing the algorithm fitted for the planet with the shorter orbital period.

part of the conservative sample of potentially habitable exoplanets (Figure 1.13).

Correctly Calculated Orbital Period Only

8 of the 50 planets found had calculated orbital period values within $\pm 3\sigma$ of the literature values listed on the NASA Exoplanet Archive but had inaccurately calculated planetary radius values.

3 of the 8 planets (Kepler-202 c, Kepler-221 b and Kepler-326 c) were from multi-planetary systems consisting of exoplanets with roughly similar sizes. The periods of the planets in these systems exhibited harmonics e.g. the Kepler-221 systems has 4 exoplanets with orbital periods of approximately 2, 5, 10 and 18 days. These overlapping periods led to a clustering of data points in the light curve (see Figure 4.4 of Kepler-221 b), resulting in t_0 and k being estimated incorrectly, hence the planetary radius value was calculated incorrectly as well. In this dissertation, we only model a single planet around each star which leads to the detection of only one planet in a multi-planetary system whose parameters are often determined incorrectly.

A further 3 of the 8 planets (Kepler-958 b, Kepler-962 b and Kepler-1114 b) were super Earths from single planetary systems with radii $< 2R_\oplus$, orbiting stars with radii between $0.8 - 0.9R_\odot$ (Morton et al. [2016]). Due to the small size of these exoplanets, the planet-star radius ratios were determined as being between $0.001 < k < 0.02$. The algorithm did not identify the transit signature. Though the transit centre was identified as one of the minimum flux values, outliers and random variations could have prevented the true transit signature from being identified (see Figure 4.5).

The remaining 2 planets of the 8 exoplanets with calculated orbital period values within $\pm 3\sigma$ of the literature values but with inaccurately calculated planetary radius values, Kepler-532 b and Kepler-533 b, have radii $> 2.5R_\oplus$ and though the transit centre has been identified correctly, there are many
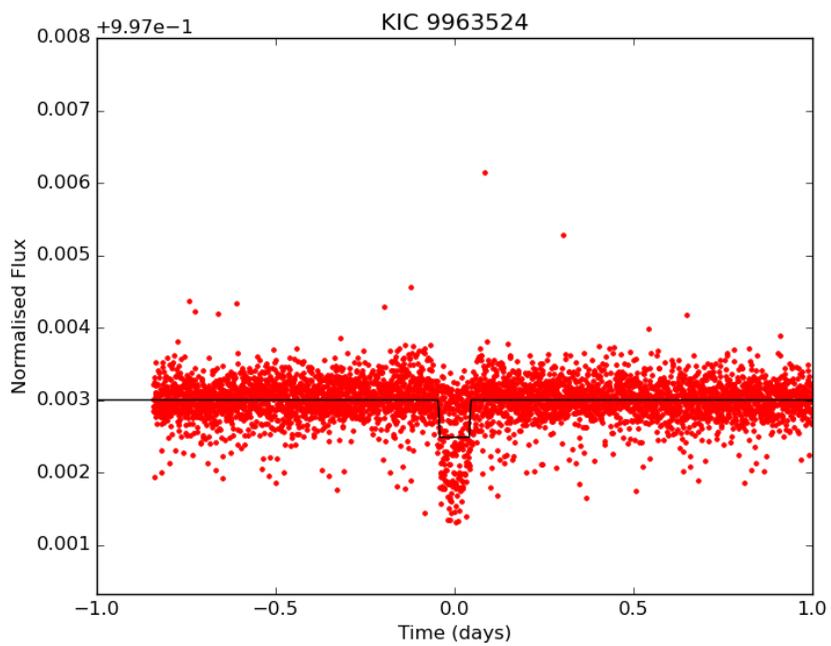
Figure 4.4: Phased transit light curve and best fit parameter model of Kepler-221 b (KIC 9963524) from 4-planet system, illustrating period harmonics.
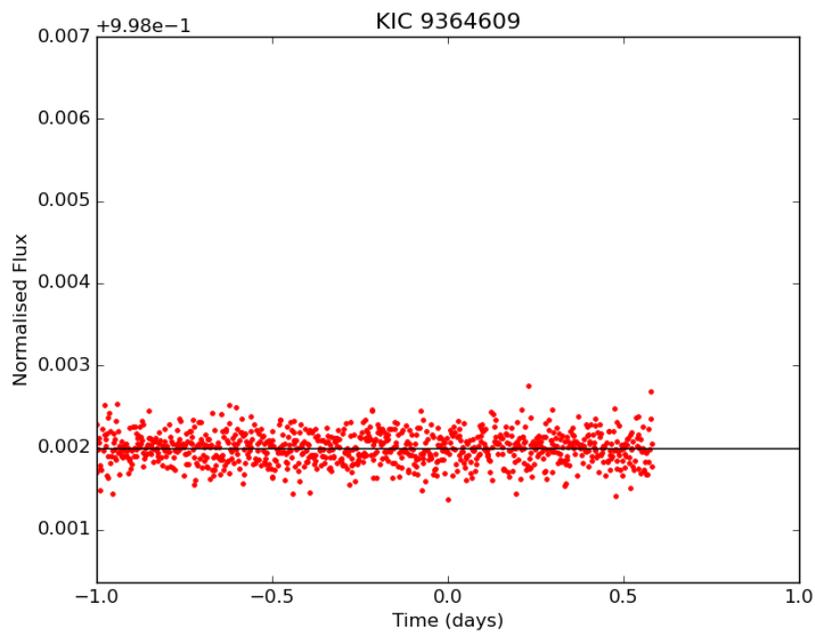
Figure 4.5: Phased transit light curve and best fit parameter model of Kepler-1114 b (KIC 9364609) with radius $< 2R_\oplus$, showing unidentified transit signal.
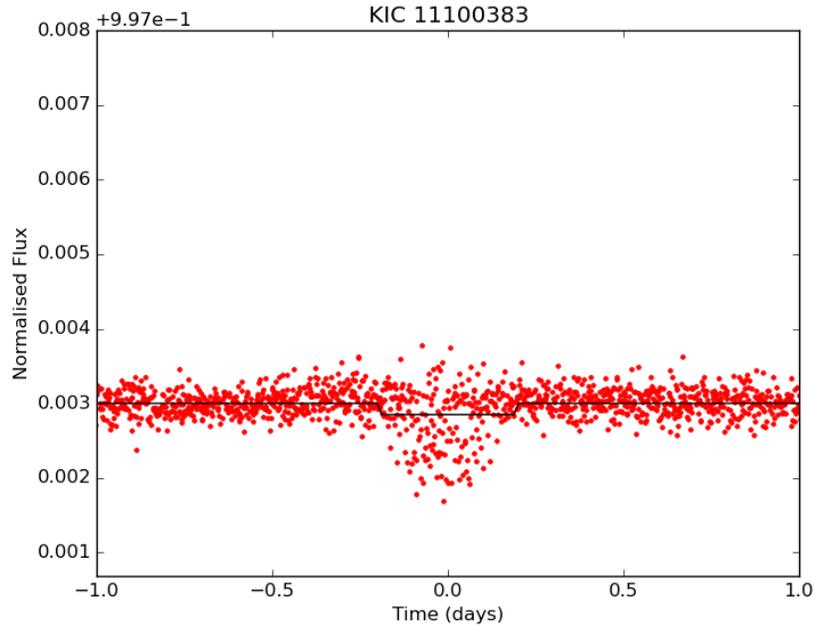
Figure 4.6: Phased transit light curve and best fit parameter model of newly validated planet Kepler-532 b (KIC 11100383), showing overlapping of data points along transit depth.

overlapping flux points along the transit depth that led to the incorrect determination of the planet-star radius ratio (see Figure 4.6 of Kepler-532 b). This resulted in the incorrect calculation of the planetary radius values. Further filtering of the data to remove random variation not due to planetary transits may help in the correct determination of k. Kepler-532 b and Kepler-533 b were newly validated as planets in 2016 in a study that also detailed 428 likely false positives. Not much follow up has been done on these planets or their stars yet (Morton et al. [2016]).

4.1.2   Shortcomings in Parameter Estimations

Multi-planetary Systems

There were 11 exoplanets out of the 50 detected planets that had incorrectly determined parameter values and came from multi-planetary systems. These planets were Kepler-9 b, Kepler-198 b, Kepler-205 b, Kepler-209 b, Kepler-215 b, Kepler-261 b, Kepler-319 b, Kepler-322 b, Kepler-755 b, Kepler-969 b and Kepler-1371 b.

There were 2 distinct trends among the light curves of these planets in multi-planetary systems:

1. The planets in the system were similar in size (and $> 1.5R_{\oplus}$) and had similar period values, leading to clustering and overlapping of data points (harmonics) for each planet (see Figure 4.7 for Kepler-261 b). This clustering of data points led to the incorrect determination of the transit centre and hence the orbital period was determined incorrectly. Since there were no distinct dips due to the overlapping of data points as well as random outliers, the planet-star radius ratio was estimated incorrectly thus the planetary radius was incorrectly determined as well. (Kepler-9 b, Kepler-198 b, Kepler-205 b, Kepler-209 b, Kepler-215 b, Kepler-261 b and Kepler-755 b).

2. The planets in the system were similar in size, but generally small ($< 1.5R_{\oplus}$), and had similar period values but the transit centre was found in a region of sparse data points (see Figure 4.8 for Kepler-322

Figure 4.7: Phased transit light curve and best fit parameter model of Kepler-261 b (KIC 2302548) from multi-planetary system showing clustering and overlapping of data points (harmonics).

b) and this was most likely due to the fact that the true transit signal was not found by the algorithm because of the small size of the planets ($0.001 < k < 0.03$). Since the transit centre was determined incorrectly, both the orbital period and planetary radius were calculated incorrectly. (Kepler-319 b, Kepler-322 b, Kepler-969 b and Kepler-1371 b).

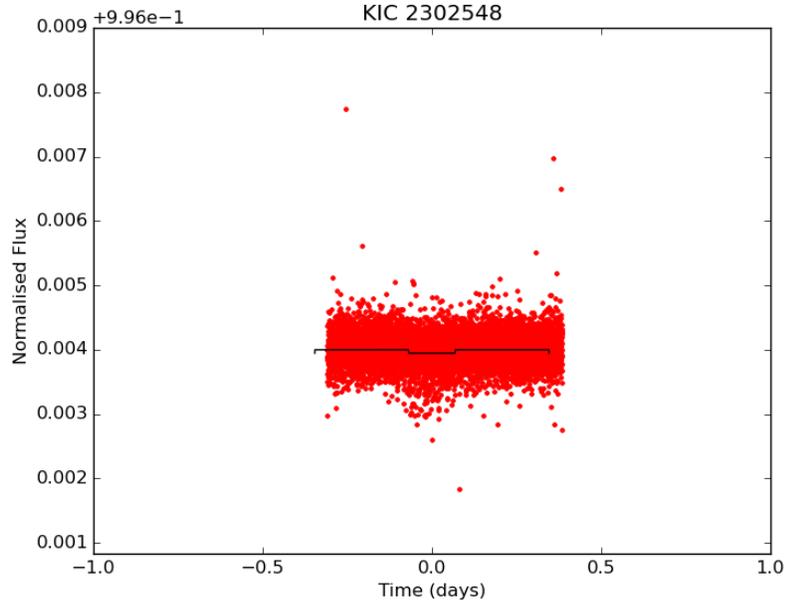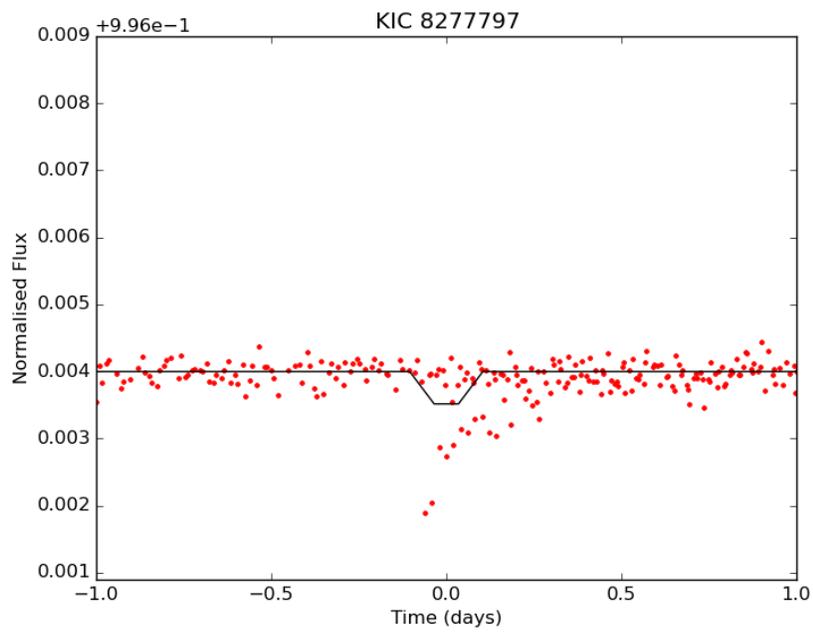Figure 4.8: Phased transit light curve and best fit parameter model of Kepler-322 b (KIC 8277797) from multi-planetary system, showing flux minimum found in a region of sparse data points.

Single-planetary Systems

There were 11 exoplanets out of the 50 detected planets that had incorrectly determined parameter values and came from single planetary systems. These planets were Kepler-479 b, Kepler-651 b, Kepler-831 b, Kepler-964 b, Kepler-984 b, Kepler-1036 b, Kepler-1178 b, Kepler-1482 b, Kepler-1499 b, Kepler-1563 b and Kepler-1558 b.

10 of these planets had radii $< 2.5R_\oplus$, with Kepler-1178 b, Kepler-1482 b and Kepler-1536 b having radii $< 1R_\oplus$, and had small planet-star radius ratios ($0.001 < k < 0.01$). Due to their small size, the algorithm did not accurately detect the true transit signal of the planets. The small planet-star radius ratio values led to almost flat models being fitted (see Figure 4.9 for Kepler-964 b). Kepler-1036 b did have a radius $> 2.5R_\oplus$ but had $k = 0.0046$ and a long orbital period (122 days) which may not have been fully captured in the 180 day observation window, resulting in incorrect calculated orbital period and planetary radius values. These issues could be resolved by using more data which will average out the noise and increase the transit signal. Having more data could also help the algorithm detect the 122 day period of Kepler-1036 b.

Period Harmonics

There were 4 exoplanets (Kepler-461 b, Kepler-477 b, Kepler-636 b and Kepler-1019 b) from single planetary systems with incorrectly determined
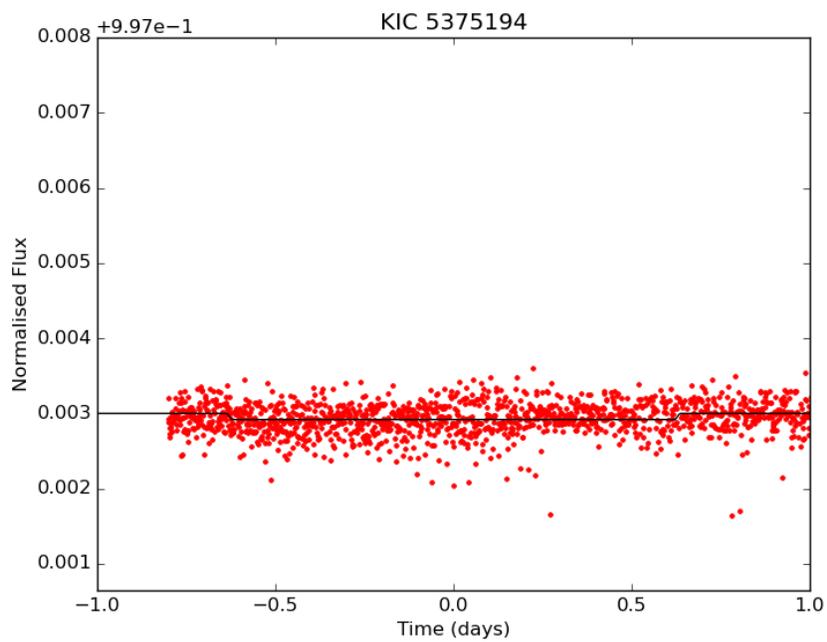
Figure 4.9: Phased transit light curve and best fit parameter model of Kepler-964 b (KIC 5375194) with radius $< 2.5R_\oplus$, showing inaccurately detected transit signal.

parameter values that exhibited orbital period harmonics i.e. the measured orbital period value was some multiple of the true (literature) orbital period value e.g. Kepler-477 b's calculated period was 5.55806 ± 0.00003 days, which is approximately $\frac{1}{2}$ the literature period value of 11.119907 ± 0.000012 days. All the planets had short orbital periods ($p < 20$ days) and a characteristic of the light curves was that the data points were aligned throughout the time series (see Figure 4.10 of Kepler-477 b).

Multi-periodic signals in time series data may exhibit harmonics of the orbital period on an exoplanet. Stellar rotation frequencies may have led to the calculated orbital period values being harmonics of the literature period values. Fourier component analysis is a prewhitening procedure that identifies the Fourier components of a time series by identifying dominant frequencies, fitting sine waves and removing the components until white noise level is achieved. Harmonic analysis can be based on the stellar rotation frequency (Hatzes [2014]). This prewhitening step may be implemented in future to remove periodic signals not due to exoplanets transits. Note: The Kepler data has already been whitened in this dissertation using our median filter.

Median Filtering

We applied a median filter to the flux values to correct for long-term trends (Fogtmann-Schulz et al. [2014]). A median filter can be used if the time-scale of the transits is shorter than the time-scale for any contribution from stellar
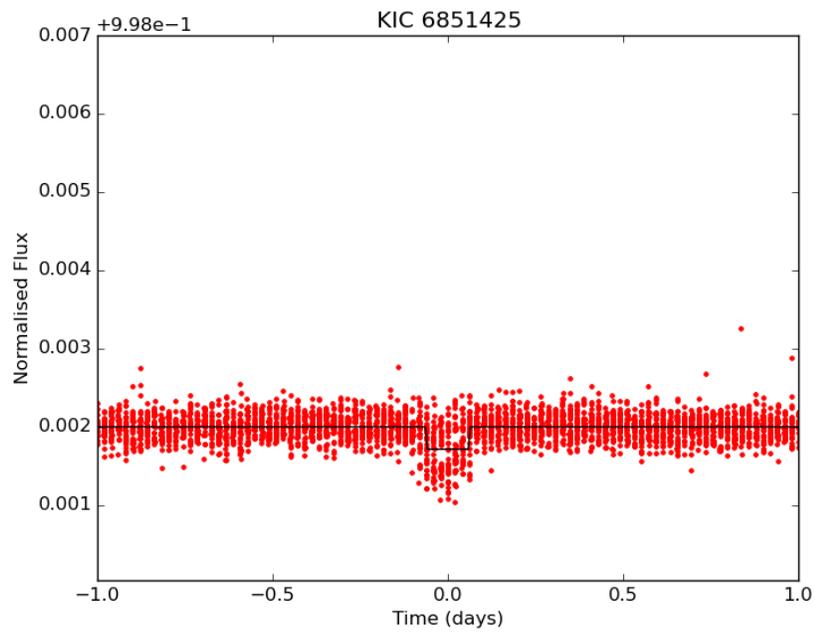
Figure 4.10: Phased transit light curve and best fit parameter model of Kepler-477 b (KIC 6851425) exhibiting orbital period harmonics.

variations in the light curves (Aigrain and Irwin [2004]). We set the median filter to remove all trends with a time scale longer than 1 day.

There were 2 exoplanets out of the 50 detected planets, Kepler-775 b and Kepler-1173 b, that had incorrectly determined parameter values because they had orbital periods of less than 1 day. Our median filter may have smoothed out the transits of these exoplanets, resulting in underestimated planet-star radius ratio values and hence incorrect calculated planetary radius values (see Figure 4.11 of Kepler-1173 b). Smoothing out the transits would also have resulted in an incorrect determination of the transit centre and thus the calculated orbital period value as well. These extremely short period planets would not be considered in the search for potentially habitable exoplanets as they certainly do not lie within the habitable zone of their host stars.

Circumbinary Planet

Kepler-453 b is a circumbinary exoplanet i.e. a planet that orbits two stars. The binary star system Kepler-453 consists of a $0.93 M_\odot$ star and a $0.194 M_\odot$ star in a 27.3 day orbit around one another (Welsh et al. [2015]). Our algorithm identified the $0.194 M_\odot$ star in the binary system and determined (correctly) the orbital period of the binary star system and the radius of the $0.194 M_\odot$ star (see Figure 4.12).

Kepler-453 b's orbit only becomes edge-on (90° inclination) in the latter part of the Kepler data i.e. quarters 9-17. Kepler-453 b has an orbital period
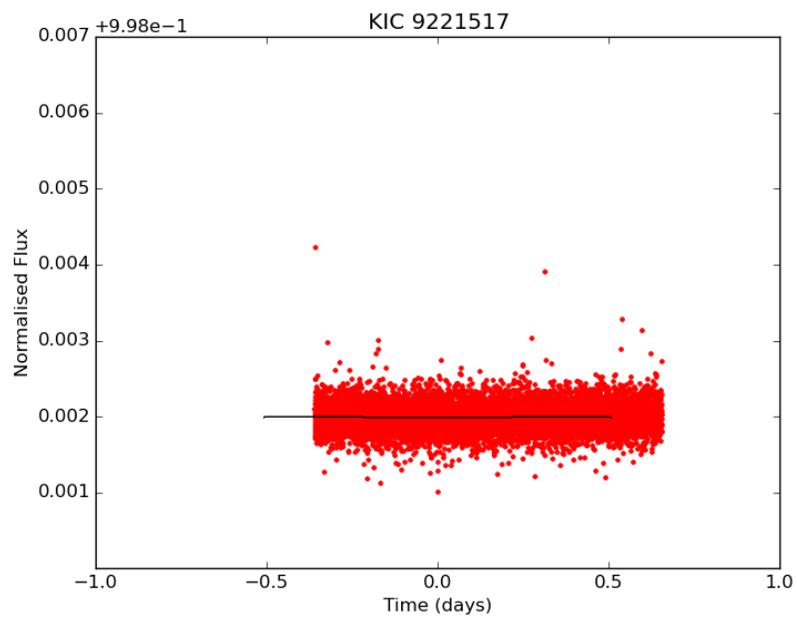
Figure 4.11: Phased transit light curve and best fit parameter model of Kepler-1173 b (KIC 9221517) with $p = 0.769$ days, affected by median filter.
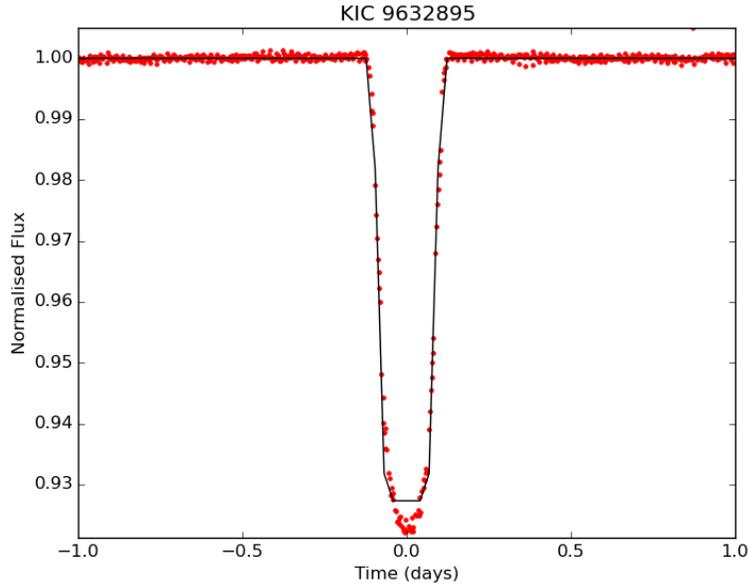
Figure 4.12: Phased transit light curve and best fit parameter model of Kepler-453 b (KIC 9632895) in binary star system.

of 240.5 days and since we used Kepler data from quarters 3 and 4 and had a maximum period threshold of 180 days, this planet went undetected by our algorithm.

Kepler-453 b was originally rejected as a false positive but once more data became available, it was re-evaluated. Kepler-453 b has a radius of $6.2R_{\oplus}$ and while it does lie within the conservative habitable zone of the binary star system, the habitable zone around a binary star is not spherical or fixed so Kepler-453 b cannot definitively be classified as a potentially habitable exoplanet (Welsh et al. [2015]).

4.2 Undetected Exoplanets

20 of the 70 known exoplanets in the blind sample went undetected by our exoplanet detection algorithm (Appendix, Table A2). The algorithm calculated the planet-star radius ratio values to be less than 0.001, and the light curves were excluded as exoplanet candidate light curves.

All the undetected exoplanets had radii $< 2.5 R_\oplus$ and orbited stars with radii between $0.7 R_\odot$ and $1 R_\odot$.

6 of the undetected exoplanets (Kepler-783 b, Kepler-1235 b, Kepler-1328 b, Kepler-1361 b, Kepler-1425 b and Kepler-1588 b) had radii $< 1 R_\oplus$ and orbited stars with radii between $0.9 R_\odot$ and $1 R_\odot$. The planet-to-star measurement signals were weak, hence the model fitted flat lines to the data as the planets made no distinctly discernible impact on the flux values of their host stars (see Figure 4.13 of Kepler-1235 b).

4 of the undetected exoplanets (Kepler-200 b, Kepler-213 b, Kepler-371 b and Kepler-390 b) were from multi-planetary systems. These planets had short orbital periods and radii $< 2.5 R_\oplus$ in planetary systems consisting of other exoplanets of similar size and orbital period, orbiting Sun-like stars $(0.95 - 1 R_\odot)$. The overlapping data points for the planets in the systems coupled with the weak planet-to-star measurement signals resulted in flat models being fitted to the data (see Figure 4.14 of Kepler-213 b) and these light curves were excluded from consideration.
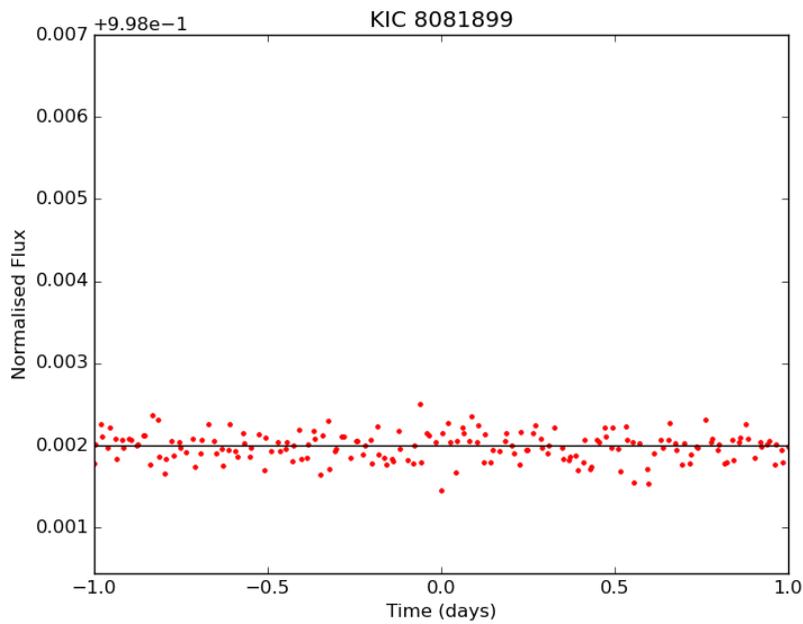
Figure 4.13: Phased transit light curve and best fit parameter model of undetected Kepler-1235 b (KIC 8081899) with $R_{planet} = 0.77R_{\oplus}$, showing weak planet-to-star measurement signal resulting in a flat model being fitted to the data.
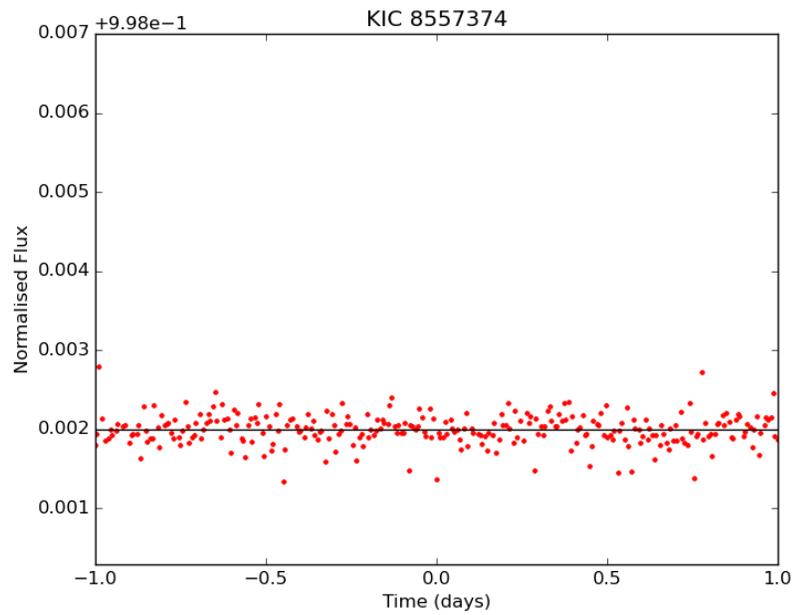
Figure 4.14: Phased transit light curve and best fit parameter model of undetected Kepler-213 b (KIC 8557374) in multi-planetary system, showing weak planet-to-star measurement signal resulting in a flat model being fitted to the data.

Kepler-922 b had a literature orbital period value of 0.938 days[2]. These exoplanet transits were smoothed out by our 1 day median filter, hence the algorithm found $k = 0.00000074$ and a flat model was fitted, resulting in Kepler-922 b going undetected.

The remaining 9 undetected exoplanets (Kepler-482 b, Kepler-640 b, Kepler-646 b, Kepler-744 b, Kepler-905 b, Kepler-906 b, Kepler-1064 b, Kepler-1240 b and Kepler-1332 b) were from single planetary systems with $1R_\oplus < R_{planet} < 2.5R_\oplus$. They had weak planet-to-star measurement signals and no distinctly discernible impact on the flux values of their host stars, which led to the fitting of flat models to the data (see Figure of Kepler-1064 b). It may be possible to increase the planet-to-star measurement signals by using more data (stitching together more light curves) in order to average out the noise in the light curves.

In order to check that these undetected planets recorded in the literature had real transit signals in the data, we plotted the transit model with its parameters set to the parameters of the exoplanets in the literature (see Figure 4.15 showing Kepler-1064 b). Even with the literature parameters passed to the model, the model did not fit the light curve data. The data could benefit from further cleaning and whitening to strengthen the planet-to-star measurement signal and the planet-star radius ratio, by further removing stellar, random and systematic variations in the light curve data.

---

[2]https://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/nph-tblView?app=ExoTbls&config=planets
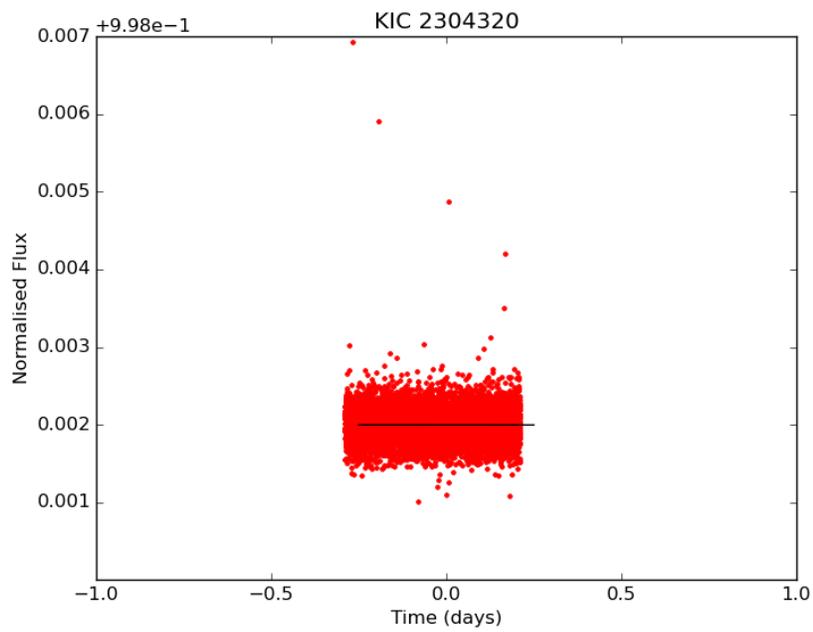
Figure 4.15: Phased transit light curve and best fit parameter model of undetected Kepler-1064 b (KIC 2304320), showing weak planet-to-star measurement signal resulting in a flat model being fitted to the data.
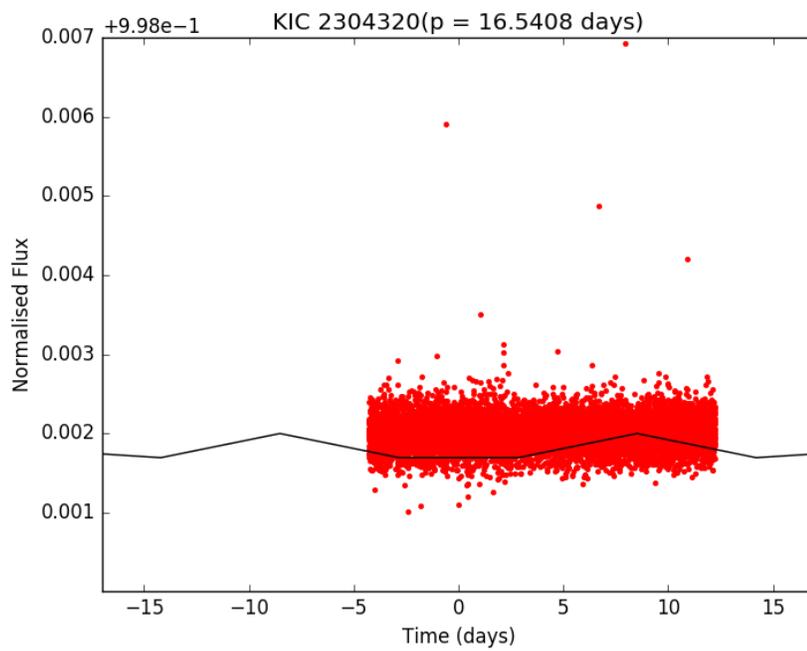
Figure 4.16: Phased transit light curve and literature parameter model of undetected Kepler-1064 b (KIC 2304320). Literature period=16.54080 days.

Chapter 5

Conclusion

We developed an exoplanet detection algorithm based on modelling transit light curves and fitting the models to light curve data using a chi-square minimization approach in order to detect exoplanets and estimate properties such as orbital period, planetary radius and semi-major axis (orbital radius) from the best fit parameters of the model.

We applied our algorithm to a blind sample of Kepler mission data consisting of approximately 4500 stars. The selection criteria for the blind sample were $T_{star} < 6000$ K, $R_{star} < 1 R_\odot$ and $13.5 <$ Kepler Magnitude < 14. The blind sample contained 70 known exoplanets.

Our algorithm detected 50 of the 70 exoplanets in the blind sample.

We found that our algorithm was effective in detecting exoplanets with planet-star radius ratios greater than 0.01 ($k > 0.01$) and/or exoplanets with radii greater than $2.5 R_\oplus$, as well as short-period exoplanets ($p < 90$ days).

There were inconsistencies in determining the semi-major axis values, a, of the exoplanets and this could be due the fact that the mass of the star (see equation (1.4)) is not available to the PyTransit package used to model the light curves. However, most of the exoplanets in the sample did not have literature semi-major axis values to compare with our calculated semi-major axis values.

There are now studies and missions looking toward small stars (K and M class stars) as opposed to Sun-like stars to search for exoplanets due to the stronger planet-to-star measurement signals (larger planet-star radius ratios), which are much more favourable for detecting Earth-sized planets (Shields et al. [2016]).

In the case of multi-planetary systems, we found our algorithm first fits for the largest transit depth and/or (subsequently) for the shortest orbital period. The algorithm finds and fits exoplanets in multi-planetary systems correctly, provided that the transits for each planet are distinct and are not overlapping, but only fits for one planet in the system. In order to find and fit more successfully for a wide range of multi-planetary systems in future, once the model has been fitted to one transiting exoplanet, we should remove the data relating to this planet from the light curve and subsequently continue fitting for the remaining planets in the system until the white noise level has been reached.

The 20 planets in the sample that went undetected were excluded by our algorithm for having planet-star radius ratios less than 0.001 ($k < 0.001$).

The planet-to-star measurement signals were weak, which resulted in no distinctly discernible impact on the flux values of the star and often flat models ($k = 0$) being fitted to the data. Exoplanet transit signals, especially of small, Earth-sized planets, can go undetected in light curves because they may be masked by random, systematic or stellar variations. Though we already performed a data whitening step (median filter), we could implement Fourier component analysis (another prewhitening procedure) that identifies the Fourier components of a time series by identifying dominant frequencies, fitting sine waves and removing the components until the white noise level is achieved. Harmonic analysis can be based on the stellar rotation frequency (Hatzes [2014]). In order to increase the planet-to-star measurement signals we could use more data in order to average out the noise in the light curves.

In this blind sample of Kepler data, no potentially habitable exoplanets were found. This is not unexpected as, of more than 3400 exoplanets found to date, only 52 exoplanets are considered potentially habitable to varying degrees i.e. 1.5% of all exoplanets found to date are considered potentially habitable. Kepler-62 d was found in a multi-planetary system with 2 long-period potentially habitable exoplanets, Kepler-62 e and Kepler-62 f, and Kepler-453 b - though found in the habitable zone of the binary star system, Kepler-453 - is not considered potentially habitable due to the uncertainty in the habitable zones of binary star systems.

The future of detecting and characterizing exoplanets in the search for an Earth twin is promising with the launch of TESS (Transiting Exoplanet

Survey Satellite) and JWST (James Webb Space Telescope) in 2018. JWST will use transit spectroscopy to observe the atmospheres of suitable rocky exoplanets found by TESS (Seager [2014]).

Appendix

KIC ID	Kepler Name	Planet-Star Radius Ratio (k)
2302548	Kepler-261 b	0.00727 ± 0.00083
2441495	Kepler-479 b	0.00923 ± 0.00086
3323887	Kepler-9 b	0.05823 ± 0.00041
3541946	Kepler-198 b	0.0078 ± 0.0007
3733628	Kepler-543b	0.02901 ± 0.00063
4180280	Kepler-472 b	0.03383 ± 0.00056
4644952	Kepler-319 b	0.0304 ± 0.0015
5113822	Kepler-199 c	0.03107 ± 0.00134
5375194	Kepler-964 b	0.00908 ± 0.00074
5796675	Kepler-636 b	0.05169 ± 0.00068
5809890	Kepler-755 b	0.00943 ± 0.00069
6020753	Kepler-202 c	0.0028 ± 0.0
6265665	Kepler-1499 b	0.00673 ± 0.00081
6362874	Kepler-775 b	0.00253 ± 0.00093
6425957	Kepler-205 b	0.0149 ± 0.0007
6851425	Kepler-477 b	0.01689 ± 0.00071
7115785	Kepler-209 b	0.01531 ± 0.00075
7211221	Kepler-831 b	0.00278 ± 0.00092
7447200	Kepler-210 c	0.04990 ± 0.00042
7626506	Kepler-112 b	0.02672 ± 0.00054
8277797	Kepler-322 b	0.02120 ± 0.0019
8359498	Kepler-77 b	0.10133 ± 0.00014
8692861	Kepler-69 b	0.0210 ± 0.0007
8802165	Kepler-648 b	0.02649 ± 0.00071
8962094	Kepler-215 b	0.0067 ± 0.0011
9002278	Kepler-62 d	0.02790 ± 0.00081
9221517	Kepler-1173 b	0.00327 ± 0.00085
9334893	Kepler-1178 b	0.0056 ± 0.0011
9364609	Kepler-1114 b	0.0010 ± 0.0
9455325	Kepler-958 b	0.00453 ± 0.00077

Table A1: Detected exoplanets in blind sample with associated star Kepler Input Catalog identification (KIC ID) and planet-star radius ratio.

KIC ID	Kepler Name	Planet-Star Radius Ratio (k)
9471268	Kepler-326 c	0.00695 ± 0.00062
9527334	Kepler-461 b	0.00917 ± 0.00055
9527915	Kepler-478 b	0.02750 ± 0.00092
9578686	Kepler-651 b	0.00816 ± 0.00092
9597058	Kepler-962 b	0.02530 ± 0.00059
9632895	Kepler-453 b	0.26956 ± 0.00022
9662811	Kepler-984 b	0.002 ± 0.001
9963524	Kepler-221 b	0.02271 ± 0.00017
10000941	Kepler-1558 b	0.00681 ± 0.00056
10190777	Kepler-1019 b	0.01191 ± 0.00055
10397751	Kepler-1371 b	0.0010 ± 0.0
10657406	Kepler-969 b	0.0120 ± 0.0015
10925104	Kepler-114 d	0.03487 ± 0.00062
11100383	Kepler-532 b	0.01229 ± 0.00092
11125797	Kepler-1482 b	0.00311 ± 0.0010
11194032	Kepler-533 b	0.03748 ± 0.00064
11359879	Kepler-15 b	0.09990 ± 0.00013
11449844	Kepler-468 b	0.14703 ± 0.00026
11769890	Kepler-1036 b	0.0046 ± 0.0
12505503	Kepler-1563 b	0.0013 ± 0.0

Table A1 continued...

KIC ID	Kepler Name	Planet-Star Radius Ratio (k)
2304320	Kepler-1064 b	0.00000041
8081899	Kepler-1235 b	0.0
5339567	Kepler-1240 b	0.0
8415200	Kepler-1328 b	0.0
8095441	Kepler-1332 b	0.00000012
9995402	Kepler-1361 b	0.0
7455981	Kepler-1425 b	0.0
10000941	Kepler-1588 b	0.0000000021
5941160	Kepler-200 b	0.0
8557374	Kepler-213 b	0.00000011
3548044	Kepler-371 b	0.0
5959719	Kepler-390 b	0.0
10810838	Kepler-482 b	0.0
6707835	Kepler-640 b	0.0
8361905	Kepler-646 b	0.00000071
10154388	Kepler-744 b	0.00000035
10337517	Kepler-783 b	0.00000018
9886661	Kepler-905 b	0.0
5009743	Kepler-906 b	0.0
11547505	Kepler-922 b	0.00000074

Table A2: Undetected exoplanets in blind sample with associated star Kepler Input Catalog identification (KIC ID) and planet-star radius ratio.

KIC ID	Kepler Name	Planet-Star Radius Ratio (k)
8191672	Kepler-5 b	0.08201 ± 0.00004
8435766	Kepler-78 b	0.00293 ± 0.00023
11804465	Kepler-12 b	0.12308 ± 0.00004
7368664	Kepler-434 b	0.05662 ± 0.00071
8866102	Kepler-410A b	0.00336 ± 0.00011
8359498	Kepler-77 b	0.10133 ± 0.00014
6922244	Kepler-8 b	0.09099 ± 0.00006
8142787	Kepler-439 b	0.0
11284772	Kepler-441 b	0.0
11757451	Kepler-443 b	0.0

Table A3: Exoplanets in test sample with associated star Kepler Input Catalog identification (KIC ID) and planet-star radius ratio.

Bibliography

S. Aigrain and M. Irwin. Practical planet prospecting. *Monthly Notices of the Royal Astronomical Society*, 350(1):331–345, 2004.

A. Baglin, M. Auvergne, P. Barge, J.-T. Buey, C. Catala, E. Michel, W. Weiss, C. Team, et al. Corot: asteroseismology and planet finding. In *Stellar structure and habitable planet finding*, volume 485, pages 17–24, 2002.

G. Bakos, R. Noyes, G. Kovács, K. Stanek, D. Sasselov, and I. Domsa. Wide-field millimagnitude photometry with the hat: A tool for extrasolar planet detection. *Publications of the Astronomical Society of the Pacific*, 116 (817):266, 2004.

B. Benneke and S. Seager. Atmospheric retrieval for super-earths: uniquely constraining the atmospheric composition with transmission spectroscopy. *The Astrophysical Journal*, 753(2):100, 2012.

D. P. Bennett. Detection of extrasolar planets by gravitational microlensing. In *Exoplanets*, pages 47–88. Springer, 2008.

A. Boccaletti, P. Baudoz, J. Baudrand, J. Reess, and D. Rouan. Imaging exoplanets with the coronagraph of jwst/miri. *Advances in Space Research*, 36(6):1099–1106, 2005.

I. A. Bond, A. Udalski, M. Jaroszyński, N. Rattenbury, B. Paczyński, I. Soszyński, L. Wyrzykowski, M. Szymański, M. Kubiak, O. Szewczyk, et al. Ogle 2003-blg-235/moa 2003-blg-53: a planetary microlensing event. *The Astrophysical Journal Letters*, 606(2):L155, 2004.

W. J. Borucki, D. Koch, G. Basri, N. Batalha, T. Brown, D. Caldwell, J. Caldwell, J. Christensen-Dalsgaard, W. D. Cochran, E. DeVore, et al. Kepler planet-detection mission: introduction and first results. *Science*, 327(5968):977–980, 2010.

T. M. Brown, D. Charbonneau, R. L. Gilliland, R. W. Noyes, and A. Burrows. Hubble space telescope time-series photometry of the transiting planet of hd 209458. *The Astrophysical Journal*, 552(2):699, 2001.

S. Csizmadia, T. Pasternacki, C. Dreyer, J. Cabrera, A. Erikson, and H. Rauer. The effect of stellar limb darkening values on the accuracy of the planet radii derived from photometric transit observations. *Astronomy & Astrophysics*, 549:A9, 2013.

T. Currie, J. Debes, T. J. Rodigas, A. Burrows, Y. Itoh, M. Fukagawa, S. J. Kenyon, M. Kuchner, and S. Matsumura. Direct imaging confirmation

and characterization of a dust-enshrouded candidate exoplanet orbiting fomalhaut. *The Astrophysical Journal Letters*, 760(2):L32, 2012.

R. Dotson, S. Seager, et al. *Exoplanets*. University of Arizona Press, 2010.

A. Fogtmann-Schulz, B. Hinrup, V. Van Eylen, J. Christensen-Dalsgaard, H. Kjeldsen, V. S. Aguirre, and B. Tingley. Accurate parameters of the oldest known rocky-exoplanet hosting system: Kepler-10 revisited. *The Astrophysical Journal*, 781(2):67, 2014.

R. L. Gilliland, W. J. Chaplin, E. W. Dunham, V. S. Argabright, W. J. Borucki, G. Basri, S. T. Bryson, D. L. Buzasi, D. A. Caldwell, Y. P. Elsworth, et al. Kepler mission stellar and instrument noise properties. *The Astrophysical Journal Supplement Series*, 197(1):6, 2011.

M. Gillon, A. H. Triaud, B.-O. Demory, E. Jehin, E. Agol, K. M. Deck, S. M. Lederer, J. De Wit, A. Burdanov, J. G. Ingalls, et al. Seven temperate terrestrial planets around the nearby ultracool dwarf star trappist-1. *Nature*, 542(7642):456–460, 2017.

A. P. Hatzes. The detection of earth-mass planets around active stars-the mass of kepler-78b. *Astronomy & Astrophysics*, 568:A84, 2014.

M. J. Holman and N. W. Murray. The use of transit timing to detect terrestrial-mass extrasolar planets. *Science*, 307(5713):1288–1291, 2005.

S. B. Howell, C. Sobeck, M. Haas, M. Still, T. Barclay, F. Mullally,

J. Troeltzsch, S. Aigrain, S. T. Bryson, D. Caldwell, et al. The k2 mission: Characterization and early results. *Publications of the Astronomical Society of the Pacific*, 126(938):398, 2014.

J. M. Jenkins, D. A. Caldwell, H. Chandrasekaran, J. D. Twicken, S. T. Bryson, E. V. Quintana, B. D. Clarke, J. Li, C. Allen, P. Tenenbaum, et al. Overview of the kepler science processing pipeline. *The Astrophysical Journal Letters*, 713(2):L87, 2010.

J. F. Kasting, D. P. Whitmire, and R. T. Reynolds. Habitable zones around main sequence stars. *Icarus*, 101(1):108–128, 1993.

J. F. Kasting, R. Kopparapu, R. M. Ramirez, and C. E. Harman. Remote life-detection criteria, habitable zone boundaries, and the frequency of earth-like planets around m and late k stars. *Proceedings of the National Academy of Sciences*, 111(35):12641–12646, 2014.

R. K. Kopparapu, R. Ramirez, J. F. Kasting, V. Eymet, T. D. Robinson, S. Mahadevan, R. C. Terrien, S. Domagal-Goldman, V. Meadows, and R. Deshpande. Habitable zones around main-sequence stars: new estimates. *The Astrophysical Journal*, 765(2):131, 2013.

K. Levenberg. A method for the solution of certain non-linear problems in least squares. *Quarterly of applied mathematics*, 2(2):164–168, 1944.

J. J. Lissauer, R. I. Dawson, and S. Tremaine. Advances in exoplanet science from kepler. *Nature*, 513(7518):336–344, 2014.

C. Lovis, F. Pepe, F. Bouchy, G. L. Curto, M. Mayor, L. Pasquini, D. Queloz, G. Rupprecht, S. Udry, and S. Zucker. The exoplanet hunter harps: unequalled accuracy and perspectives toward 1 cm s- 1 precision. In *Proc. SPIE*, volume 6269, pages 62690P1–62690P23, 2006.

K. Mandel and E. Agol. Analytic light curves for planetary transit searches. *The Astrophysical Journal Letters*, 580(2):L171, 2002.

D. W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, 11(2):431–441, 1963.

J. Mason. *Exoplanets: detection, formation, properties, habitability*. Springer Science & Business Media, 2008.

M. Mayor and D. Queloz. A jupiter-mass companion to a solar-type star. *Nature*, 378, 1995.

M. Mayor, C. Lovis, and N. C. Santos. Doppler spectroscopy as a path to the detection of earth-like planets. *Nature*, 513(7518):328–335, 2014.

C. P. McKay. Requirements and limits for life in the context of exoplanets. *Proceedings of the National Academy of Sciences*, 111(35):12628–12633, 2014.

T. D. Morton, S. T. Bryson, J. L. Coughlin, J. F. Rowe, G. Ravichandran, E. A. Petigura, M. R. Haas, and N. M. Batalha. False positive probabilities

for all kepler objects of interest: 1284 newly validated planets and 428 likely false positives. *The Astrophysical Journal*, 822(2):86, 2016.

M. Newville, T. Stensitzki, D. B. Allen, and A. Ingargiola. LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for PythonÂű, Sept. 2014. URL `https://doi.org/10.5281/zenodo.11813`.

H. Parviainen. PYTRANSIT: fast and easy exoplanet transit modelling in PYTHON. *MNRAS*, 450(April): 3233–3238, 2015. doi: 10.1093/mnras/stv894. URL `http://mnras.oxfordjournals.org/cgi/doi/10.1093/mnras/stv894`.

J. Pepper, R. W. Pogge, D. DePoy, J. Marshall, K. Stanek, A. M. Stutz, S. Poindexter, R. Siverd, T. P. OâĂŹBrien, M. Trueblood, et al. The kilodegree extremely little telescope (kelt): A small robotic telescope for large-area synoptic surveys. *Publications of the Astronomical Society of the Pacific*, 119(858):923, 2007.

J. Pepper, R. B. Kuhn, R. Siverd, D. James, and K. Stassun. The kelt-south telescope. *Publications of the Astronomical Society of the Pacific*, 124(913):230, 2012.

D. Pollacco, I. Skillen, A. C. Cameron, D. J. Christian, C. Hellier, J. Irwin, T. Lister, R. Street, R. G. West, D. Anderson, et al. The wasp project and the superwasp cameras. *Publications of the Astronomical Society of the Pacific*, 118(848):1407, 2006.

M. J. Powell. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *The computer journal*, 7 (2):155–162, 1964.

W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, et al. *Numerical recipes*, volume 3. cambridge University Press, cambridge, 1989.

I. Ridpath. *A dictionary of astronomy*. Oxford University Press, 2012.

J. L. Russell. Kepler's laws of planetary motion: 1609–1666. *The British Journal for the History of Science*, 2(1):1–24, 1964.

N. C. Santos. Extra-solar planets: Detection methods and results. *New Astronomy Reviews*, 52(2):154–166, 2008.

E. Schlawin, M. Zhao, J. K. Teske, and T. Herter. A 0.8-2.4 μm transmission spectrum of the hot jupiter corot-1b. *The Astrophysical Journal*, 783(1): 5, 2014.

S. Seager. Exoplanet habitability. *Science*, 340(6132):577–581, 2013.

S. Seager. The future of spectroscopic life detection on exoplanets. *Proceedings of the National Academy of Sciences*, 111(35):12634–12640, 2014.

S. Seager and W. Bains. The search for signs of life on exoplanets at the interface of chemistry and planetary science. *Science advances*, 1(2):e1500047, 2015.

S. Seager and D. Deming. Exoplanet atmospheres. *Annual Review of Astronomy and Astrophysics*, 48:631–672, 2010.

A. L. Shields, S. Ballard, and J. A. Johnson. The habitability of planets orbiting m-dwarf stars. *Physics Reports*, 663:1–38, 2016.

S. Thompson, D. Fraquelli, J. Van Cleve, and D. Caldwell. Kepler archive manual. Technical report, KDMC-10008-004, 2012.

J. Trauger, K. Stapelfeldt, W. Traub, J. Krist, D. Moody, D. Mawet, E. Serabyn, C. Henry, P. Brugarolas, J. Alexander, et al. Access–a concept study for the direct imaging and spectroscopy of exoplanetary systems. Society of Photo-Optical Instrumentation Engineers (SPIE), 2010.

W. F. Welsh, J. A. Orosz, D. R. Short, W. D. Cochran, M. Endl, E. Brugamyer, N. Haghighipour, L. A. Buchhave, L. R. Doyle, D. C. Fabrycky, et al. Kepler 453 b âĂŤ the 10th kepler transiting circumbinary planet. *The Astrophysical Journal*, 809(1):26, 2015.

A. Wolszczan and D. A. Frail. A planetary system around the millisecond pulsar psr1257 (plus) 12. *Nature*, 355(6356):145, 1992.

J. Yang, G. Boué, D. C. Fabrycky, and D. S. Abbot. Strong dependence of the inner edge of the habitable zone on planetary rotation rate. *The Astrophysical Journal Letters*, 787(1):L2, 2014.

T. Yaqoob. *Exoplanets and alien solar systems*. New Earth Labs, 2011.