Understanding Exoplanets and Other Variable Sources in Sparsely-Sampled Time Domain Surveys

By

Michael B. Lund

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Physics

May, 2017

Nashville, Tennessee

Approved:

Keivan G. Stassun, Ph.D.

Joshua Pepper, Ph.D.

David A. Weintraub, Ph.D.

Andreas A. Berlind, Ph.D.

Paul Sheldon, Ph.D.

DEDICATION

It is said that it takes a village to raise a child. It could just as easily be said that it takes a galaxy to form a star, and I have been very fortunate to have had a stellar association of people that have helped me along the way.

First of all, I want to thank my family, starting with the constant support of my mom and dad (Kathi and Bryant) and my brother (Tom), as well as my grandparents, Grandpa and Grandma Lund and Grandma Fred. I'm also very appreciative of the financial help I've gotten from my Uncle Russ and Aunt Julie, and the support that I've received from my entire extended family while spending over a decade in college.

I owe a great deal to the members of the Local Group Astronomy Club of the Santa Clarita Valley (and its founder Laverne Booth), who first welcomed me in as a four-year-old, introduced me to so much about astronomy and taught me the night sky, and always treated me as a friend rather than talking down to me because of my age. I hope that I can always carry that attitude into any manner of outreach or mentoring that I do in the future.

I have had the benefit of a number of great friends that have been supportive over the years that have helped me keep going, and so I thank those that have stuck with me since high school, everyone that I've met lining up for movies, all my colleagues and coworkers at Griffith Observatory, and all the other assorted people that I've managed to pick up somewhere along the way. There's no greater force to keep going in astronomy than a large group of supportive nerds for friends. As the last five years have been a particularly rough stretch for me personally, thanks to Caroline Ritter, Sean Dennehy, Missy Coffman, Liz McPartland, Dani Satow, Kim Wroe, John Franz, Erica Fey, and anyone I may have foolishly omitted for help through those rough patches, as they were proof that not all dark time is good for an astronomer. I'm particularly lucky for all the love and support that I've gotten from Katherine Lee during the final stretch.

I would not have gotten to Vanderbilt without the support of many people throughout

my education. This includes the Physics and Astronomy department at California State University, Northridge, my Master's adviser Damian Christian (who got me first working with exoplanets), the former chair Christina Cadavid, the rest of the faculty, and my fellow students.

I greatly appreciate Keivan G. Stassun and Joshua Pepper giving me a chance at a Ph.D. at Vanderbilt, the research and networking opportunities that I've had over the last five years thanks to them, and pushing me develop my professional and academic skill sets. I would like to thank rest of the members of my dissertation committee, David A. Weintraub, Andreas A. Berlind, and Paul Sheldon for their time, advice, and expertise. Thanks also to the rest of the Vanderbilt astronomy community that have made it such a welcoming place, particularly my frequent sounding boards and/or collaborators, Joey Rodriguez, Kyle Conroy, Savannah Jacklin, and Rob Siverd. Finally, while it frequently goes unacknowledged, graduate school can carry a particularly heavy burden on mental health, and I would not have made it this far without the resources provided by Vanderbilt's Psychological & Counseling Center.

ACKNOWLEDGMENTS

We acknowledge support through NASA ADAP grant NNX12AE22G and through the Vanderbilt Initiative in Data-intensive Astrophysics.

We thank the attendees of the LSST Observing Cadences Workshop, and the fast transient group led by Mansi Kasliwal and Lucianne Walkowicz in the early development of the Observation Triplets metric. We also are grateful for the assistance from the LSST team in troubleshooting issues with the MAF.

The authors would like to thank Kyle Conroy for multiple helpful discussions on this research. The DASCH project at Harvard is grateful for partial support from NSF grants AST-0407380, AST-0909073, and AST-1313370.

For the development of the microlensing metric, we are grateful for the input from Rachel Street, Radek Poleski, David Ciardi, and Will Clarkson.

This work has made use of NASA's Astrophysics Data System and the SIMBAD database operated at CDS, Strasbourg, France.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 Time Domain Surveys

Astronomy has greatly benefited from the rate at which technology has been developing in the last few decades. High-quality CCDs and increasing efficiency in digital storage and processing have helped to usher in the era of big data. For observational astronomy, one of the major impacts that this has had is the proliferation of time domain surveys, and the number of objects and observations that each of these surveys can be comprised of. The scope of these surveys will continue to increase as the technological limitations are pushed back, and it is becoming increasingly important to understand the capability of these surveys.

One of the largest such surveys currently under development is the Large Synoptic Survey Telescope (LSST), which is currently under construction with a ten-year mission beginning in 2020. LSST will produce light curves for approximately 1 billion objects in a survey covering half the sky. LSST also brings with it an additional challenge; as LSST will attempt to survey such a large area of sky, the light curves it generates will have a very low cadence, particularly compared to most surveys that have been used previously for understanding transient and variable sources. The lessons that can be learned with LSST to find variables despite the low cadence will be directly applicable to other high-precision sparsely-sampled data sets from other time domain surveys and play a key role in maximizing the science output in the big data era.

1.2 Exoplanets

One of the science cases that has benefited greatly from the move toward big data in astronomy is the search for exoplanets. The first exoplanets to be discovered came in at a

slow rate as the field was pioneered in the 1990s and early 2000s, with the first confirmed planets in 1992 [Wolszczan and Frail, 1992], the first planet around a main sequence star in 1995 [Mayor and Queloz, 1995], and the first planet discovered through the transiting method in 2003 [Konacki et al., 2003]. There are now almost 3,000 confirmed exoplanets and the *Kepler* mission alone has provided another 2,500 planet candidates. However, these planets have primarily been limited to relatively close to our location in the Milky Way (Figure 1.1), with a large portion occurring in the direction of the *Kepler* field, as shown in Figure 1.2.

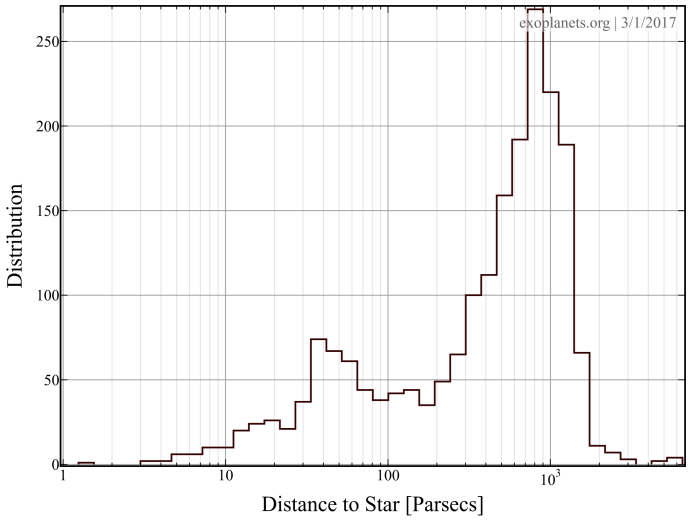

Figure 1.1: A histogram generated by exoplanet.org depicting the distances to exoplanet host stars. A 'distant' peak reflects the distance to the *Kepler* field and the closer peak represents the combined efforts of most other planet searches.

The limited populations of stars that have been searched for exoplanets has meant that while the frequency of exoplanets is being constrained particularly for solar-type field stars, many other stellar populations have not been as thoroughly examined. Red and white dwarfs, stars in clusters, stars in the Galactic core, and stars in the Magellanic Clouds have been comparatively unexplored, and understanding the planet frequencies for these stellar populations will help shed light on the mechanics for exoplanet formation.

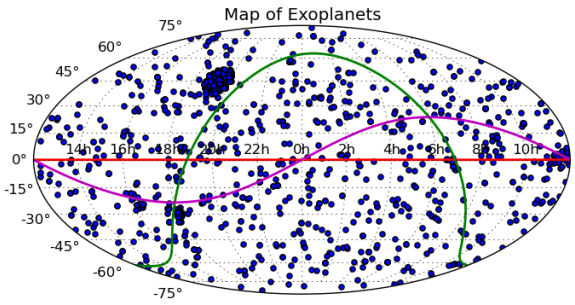

Figure 1.2: The distribution of confirmed exoplanets throughout the sky, showing that a large portion are from the *Kepler* field specifically, and that the full distribution of stars in the galaxy is not reflected in the exoplanet distribution. Figure first released in Lund [2016]

Chapter 2

Transiting Planets with LSST. I. Potential for LSST Exoplanet Detection

2.1 Introduction

The ongoing search for exoplanets has to date yielded a sample of some 1,500 planets, primarily from the contributions of the *Kepler* mission [Batalha et al., 2013], ground-based transit surveys, such as HATnet [Bakos et al., 2011], TrES [O'Donovan and Charbonneau, 2007], SuperWASP [Doyle et al., 2013], XO [Poleski et al., 2010], KELT [Pepper et al., 2007], and RV surveys, such as the Anglo-Australian Planet Search [Tinney et al., 2001], HARPS [Zechmeister et al., 2013], California and Carnegie Planet Search [Patel et al., 2007], and the Lick Planet Search [Fischer et al., 2014]. However the current sample predominantly comprises host stars that are F, G, and K dwarfs in the solar neighborhood. Due to a variety of observational and selection biases, other populations have not been examined as thoroughly, including late-type stars, evolved stars, stars outside the solar neighborhood, young stars, stellar remnants, and others. As a consequence, while the current sample of planets has allowed us to characterize exoplanet frequency as a function of mass, radius, and period [Howard et al., 2012, Petigura et al., 2013, Youdin, 2011], the current exoplanet sample does not include stellar hosts that are fully representative of all possiblities.

In this paper[1], we introduce an analytic- and simulation-based effort to explore the ability of the upcoming Large Synoptic Survey Telescope [LSST Science Collaboration et al., 2009] to discover transiting exoplanets, specifically among distant stellar populations and late-type stars. Since LSST was not specifically designed for this purpose, there are at least two important difficulties involved. The two most prominent are: (1) The stars observed by LSST are too faint for candidates to be confirmed as exoplanets through standard ob-

[1] This chapter was originally published as [Lund et al., 2015]

servational methods, making it difficult to separate real exoplanets from the many types of false positives, and (2) the cadence of LSST observations is much lower than current transit surveys, making robust detection of transit-like events against background noise more difficult. On the other hand, LSST will observe exponentially more stars than prior surveys, and it is prudent to examine just what kinds of exoplanets can be identified. Even if it proves impractical to comfirm most LSST-identified exoplanet candidates via traditional followup observations, it may be possible to derive bulk statistical properties for this exoplanet population.

This paper is the first in a series that will model the predicted properties of the LSST data set to examine:

- What tools can be used to extract exoplanet transit candidates from the LSST light curves?
- What will be the dominant types of false positives?
- How can the likelihood of detecting these false positives be mitigated, using either the LSST data themselves or follow-up observations?
- What new populations of stars can be probed for planets with LSST that have not yet been searched thoroughly?
- What do we predict will be the number and parameter distribution of exoplanets detectable with the LSST?

This sort of inquiry has been partially addressed by other papers. The LSST Science Book (LSST Science Collaboration et al. [2009], see Section 8.11), and separately Beatty and Gaudi [2008], examine how many transiting planets might be detected by LSST. We intend to build on these initial examinations by utilizing detailed LSST cadence and photometric performance attributes applied to different types of stars, as well as incorporating updated information about the intrinsic distribution of planet sizes and periods. In later papers will will also include realistic models of stellar binarity and variability, as well as other false positive sources, and investigate improvements to existing transit-detection algorithms

customized for the LSST cadence and multiband data.

In this first paper, we present our statistical framework, and we demonstrate that certain types of interesting transiting planets should be detectable in LSST data. We describe our method for modeling LSST light curves in §2.2 and describe simple models for simulating and recovering transiting exoplanets in the light curves (§2.3). In §2.4 we present the results of this study, in which we demonstrate that a number of scientifically interesting planet/star configurations should be readily detectable in LSST data, even with the use of standard transit-search routines. We close by discussing the key challenges of this project in §2.5.

2.2 Assumed LSST Parameters

The Large Synoptic Survey Telescope (LSST) is an 8.4-meter telescope being constructed in Cerro Pachon, Chile, with first light planned in 2020 [Ivezic et al., 2008]. LSST will survey approximately 30,000 square degrees of the sky repeatedly over the course of 10 years. Observation of a field will consist of two consecutive exposures of 15 seconds each, hereafter refered to as 'visits'. For the majority of the sky, each field will have $\sim$ 1000 visits. This will result in approximately 1 billion light curves that are sparsely sampled over time. LSST will also have a few selected fields that will be observed at a much higher frequency, referred to as deep-drilling fields. Each of these fields will represent 1% of the total number of LSST observations and will receive $\sim$ 10,000 visits over LSST's scheduled operation. For both cadences (regular cadence and deep-drilling), these observations will be distributed among 6 different wavelength bands, *ugrizy*. LSST's sensitivity will cover a magnitude range from $m \sim 16$ down to $m \sim 25$. In this section we describe the cadence and noise parameters that we adopt in our simulated light curves.

2.2.1 Cadence Model

The LSST team has developed the LSST Operations Simulator (OPSIM) software tool that includes models of the hardware and software performance, site conditions, and ca-

dence as a function of sky position [Delgado et al., 2014]. For our purposes, OPSIM produces a list of observation times with field ID, band, and limiting magnitude. For this paper we use the OPSIM v2.3.2, run 3.61 results[2] for two fields — one representative of a standard field, and one representative of a deep-drilling field — in order to examine the effect of the two LSST cadences on the observed lightcurves.

In Figure 2.1, we show the time between consecutive exposures for both fields. In both cases, we have excluded nine intervals that represent the gaps between seasons of observing, which are on the order of 200 days. For the deep drilling field, it is worth noting that because the observing schedules focus on hour-long blocks of constant observation, there is a large peak at $\sim$ 35 seconds representing exposure time plus readout time for consecutive observations of the same field. This then represents the minimum interval between most observations.

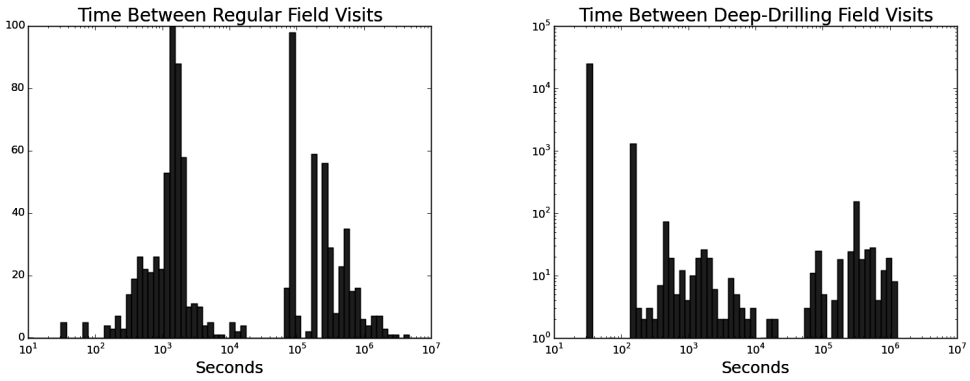

Figure 2.1: Time between beginning of subsequent exposures for a given field, for both regular cadence (left) and deep-drilling cadence (right). Note that a large fraction of the deep-drilling observations are taken in immediate succession, resulting in the largest peak at $\sim$ 34 seconds, corresponding to the combined LSST exposure and readout times.

2.2.2 Noise Model

The photometric noise in simulated LSST light curve comprises two components, a systematic noise floor and random photon noise. LSST's systems are designed to maintain

[2] Available at https://www.lsstcorp.org/?q=opsim/home

low systematic error with $\sigma_{sys} < 0.005$ mag. For purposes of this paper, the systematic error is treated conservatively as $\sigma_{sys} = 0.005$ mag. The random photon noise, dependent on the band being used, varies with changes in atmospheric extinction, stellar magnitude, and seeing conditions. Ivezic et al. [2008] define the random photometric error with the following equation:

$$\sigma^2_{rand} = (0.04 - \gamma)x + \gamma x^2 \qquad [\text{mag}^2] \quad (2.1)$$

where γ is a band-specific parameter, and $x = 10^{(m-m_5)}$. Here, m is the band-specific apparent magnitude and m_5 is the 5σ limiting magnitude for point sources. The quantity m_5 includes both band-specific parameters and factors such as seeing conditions, air mass, and exposure time. We can then express the total noise for a single visit:

$$\sigma^2 = \sigma^2_{sys} + \sigma^2_{rand} \qquad [\text{mag}^2] \quad (2.2)$$

Figure 2.2 shows the total noise for a single LSST visit to any field at the zenith, adopting m_5 and γ from Ivezic et al. [2008]. As transits of large exoplanets frequently have depths on the order of 0.01 mag, Figure 2.2 shows that for apparent magnitude ~ 16 down to as far as ~ 20, for some bands the photometric noise is less than the depth that would be expected of a transit event. This represents a much fainter apparent magnitude range with this level of photometric precision than most previous transiting planet searches have explored.

2.3 Simulated LSST Light Curves of Transits

2.3.1 Light Curve Construction

We create synthetic light curves to analyze the ability to detect transit signatures. In §2.4 we define several star/planet configurations for simulated light curves, specifying for each the stellar mass, planetary radius, orbital period, and distance from Earth. Stellar

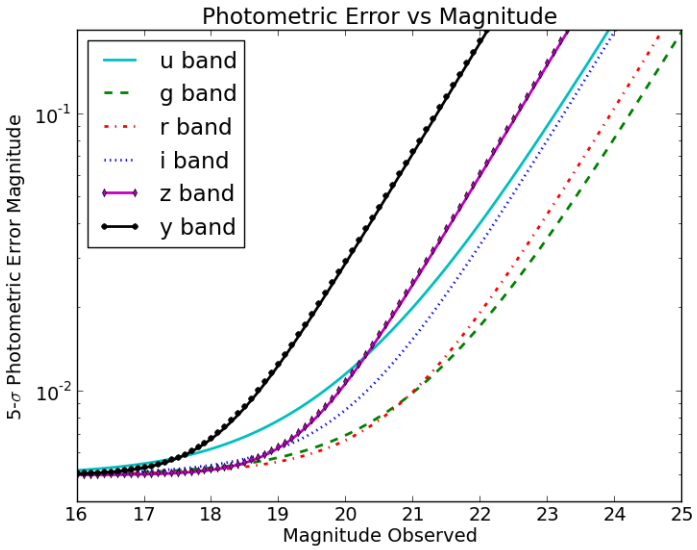

Figure 2.2: Noise model at zenith for LSST as a function of stellar magnitude in all observed bands. From Ivezic et al. [2008]

masses are converted to spectral types by interpolating the relations from Allen and Cox [2000]. Based on spectral type, we obtain absolute magnitudes in the *ugriz* bands from Covey et al. [2007]. While there is still much uncertainty about the specific parameters of the *y*-band filter that LSST will eventually use, for this work we use a *y*-band defined by Hodgkin et al. [2009].

Light curve generation itself requires orbital period, as well as transit depth and duration as determined by the geometry of the system, and the apparent magnitude of the host star in all bands. *Kepler* has shown that while stars do have a wide range of intrinsic stellar variability, most will have variability of less than 1% [Basri et al., 2011]. This intrinsic noise will be smaller than both the noise present in the observations and the transit depths we focus on, so we do not include that effect in these light curves. Transits are modeled as boxcars and when applied to the apparent magnitude of the star, allow us to create a noiseless light curve of the star that features the transit. We then add the photometric noise to the light curve, modeled as Gaussian noise, and dependent on apparent magnitude and band (See Figure 2.2) to create light curves similar to what LSST will observe.

The resulting light curve can be represented in all six bands. Limb darkening is a second-order effect that will tend to decrease the transit depth at blue wavelengths relative to a simple box-shaped transit. However, to first order the transit depth will be independent of wavelength, so in this initial exploration we treat the transits as achromatic boxcar events. With this simplification we can median-subtract each light curve and combine them to create one unified light curve for enhanced transit detection. An example of a six-band light curve is shown in Figure 2.3. It is already noticeable in this case that the *u*-band will have the largest noise for the stellar populations in which we are most interested.

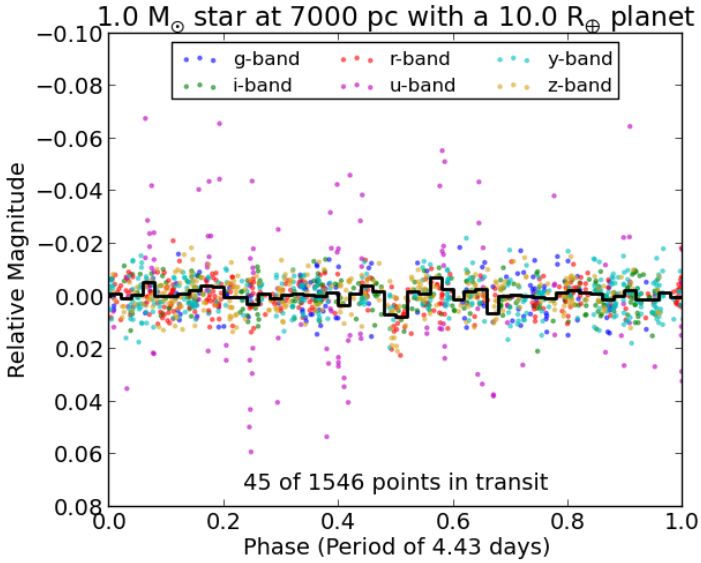

Figure 2.3: A sample six-band light curve, demonstrating the ability to combine all bands into a single phased light curve. The black line shows the binned data for the light curve.

The multiband nature of the observations will also provide additional information to characterize the light curves. Since exoplanet transits would be expected to be the same depth in all wavelengths, signals from eclipsing binary stars can be eliminated as planet candidates if we observe transit depths that are dependent on wavelength. Our intention in this paper is not to provide a rigorously developed number for expected exoplanet detections, or to address distinguishing planetary transits from other signals, but rather to demonstrate that exoplanets will be theoretically detectable, and that this is a worthwhile

consideration as the LSST project moves forward.

2.3.2 Period Recovery

We examine the recoverability of transits by searching for periodicity using the Box-fitting Least Squares (BLS) algorithm [Kovacs et al., 2002][3]. The BLS algorithm has been shown to be extremely powerful for detecting transits in high cadence data sets [Enoch et al., 2012]. While LSST will not be a high cadence survey, we use this algorithm as an initial test of LSSTs capability for recovering transit signals. In order to examine the significance of these detections, we also measure the liklihood of having a peak of equal strength in the BLS periodogram by random chance, or due to some feature of the cadence, to measure a false alarm probability (FAP). To carry out this check, for each light curve we randomly rearrange the magnitude and time information, and then reapply BLS. We conduct that operation 1000 times for each light curve, and record the highest BLS peak in each case. The resulting data sets allows us to specify the 1% and 0.1% false alarm probabilities for BLS.

2.4 Results: Examples of Simulated Light Curves for Different Types of Transits

We begin examining detectability by creating simulated light curves (§2.3) to examine what exoplanet transits will look like for several different cases with stellar parameters outlined in Table 2.1. These are intended to exemplify well-studied populations of exoplanets as well as some populations that could prove very interesting but have not been a focus of planet searches thus far.

[3] We use the implementation of the BLS algorithm that is part of the VARTOOLS package[Hartman et al., 2008]

Table 2.1: Photometric Properties of Example Systems

Scenario		Band					
		m_u	m_g	m_r	m_i	m_z	m_y
G-dwarf at 7000 pc	Apparent Magnitude	21.433	20.066	18.699	18.540	18.500	17.913
	Standard Deviation	0.023	0.007	0.006	0.006	0.006	0.007
K-dwarf at 2000 pc	Apparent Magnitude	23.619	21.091	18.526	17.980	17.675	17.265
	Standard Deviation	0.149	0.011	0.005	0.005	0.005	0.006
M-dwarf at 400 pc	Apparent Magnitude	24.424	21.466	18.530	17.150	16.428	15.704
	Standard Deviation	0.286	0.013	0.005	0.005	0.005	0.005
G-dwarf at 50 kpc	Apparent Magnitude	25.658	24.334	22.968	22.807	22.770	22.183
	Standard Deviation	1.029	0.109	0.043	0.068	0.122	0.214

Absolute magnitudes for *ugriz* bands from Covey et al. [2007] and *y* band from Hodgkin et al. [2009], with corresponding photometric noise.

2.4.1 A Hot Jupiter Orbiting a G-dwarf in the Milky Way

The category of transiting exoplanets that have been studied most extensively are the Hot Jupiters, Jovian-mass planets orbiting sun-like stars with periods on the order of days. Here, we simulate a light curve for a 10.0 $R_\oplus$ planet orbiting at 4.43 days around a 1.0 $M_\odot$ star. We have set the star at a distance of 7000 pc, chosen so that the apparent magnitude of the host star is within the sensitivity range of LSST in all bands, while also representing a distance inaccessible to most current exoplanet surveys. The resultant light curve is then phase-folded on the input period in order to generate the phased light curve displayed in Figure 2.4. In this case, there is a very noticeable drop in the star's brightness of 0.009 mag, showing that LSST can at least detect such a transit strong enough in the six-band lightcurve to be visually identifiable.

For the Hot Jupiter case, we provide the BLS periodograms for both a regular LSST field and a deep-drilling field in Figure 2.5. We also mark the true period on the periodogram, as well as the aliases at half and double the period. In both cases, the 4.43 day period is the highest peak in the periodogram, providing a qualitative reassurance that that the periods visible in Figure 2.4 could be recovered had the value not been known. We

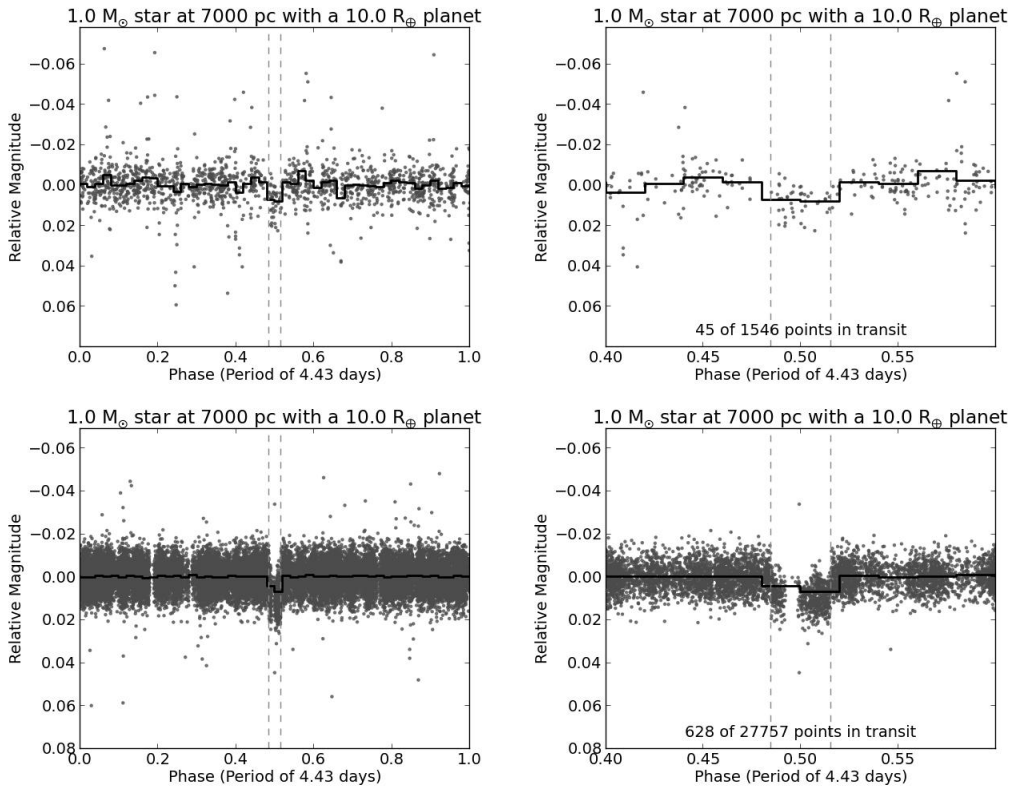

Figure 2.4: Light curve for a 10.0 $R_\oplus$ planet in a 4.43 day period around a 1.0 $M_\odot$ star at 7000 pc. The top two plots show a regular LSST field and the bottom two plots show an LSST deep-drilling field. The plots on the left show the full phase, and the plots on the right zoom in on the transit. Black lines are binned data of the light curve.

find that in 1000 light curves there are no peaks of equal or greater height, from which we determine the false positive rates for Hot Jupiters at both cadences to be <0.1%.

2.4.2 A Hot Neptune Orbiting a K-dwarf in the Milky Way

Next we examine a Hot Neptune. Representing a regime between very large Jovian planets and smaller terrestrial planets (radii 3-4 $R_\oplus$) that has been richly populated by *Kepler*. Here we place a 4.0 $R_\oplus$ planet in a 7.3 day period around a 0.6 $M_\odot$ star at 2000 pc.

Ground-based observations are generally limited in sensitivity to transit depths on the order of 1%, so we have used a slightly smaller host star such that a Neptune-sized planet

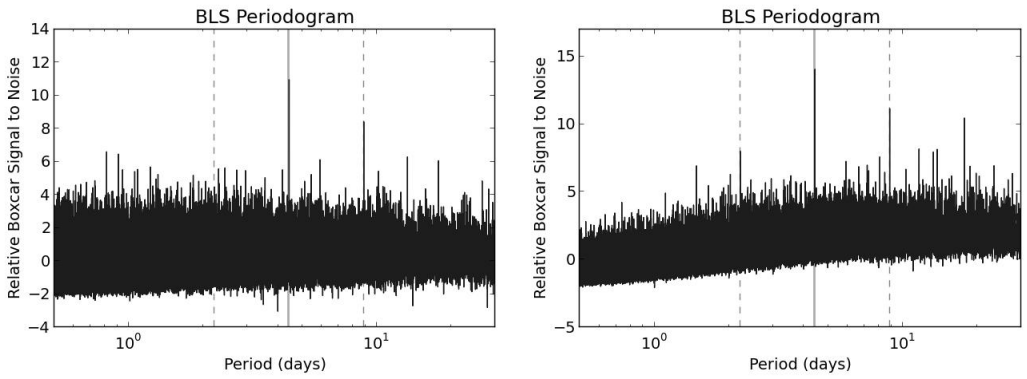

Figure 2.5: BLS periodogram for a Hot Jupiter in a standard field on the left, and a deep drilling field on the right. The lighter grey line marks the actual period of 4.43 days, while the other two dashed lines mark twice and half the true period.

will still represent an approximately similar drop in brightness from the Hot Jupiter case. The closer distance is to maintain the constraint that the star's apparent magnitude in all bands will be within the sensitivity range of LSST. As the transit events that we are examining will generally range between 3 mmag and 10 mmag in depth, we chose only to plot bands where the standard deviation is less than 30 mmag. In this light curve the noise in the *u*-band dominates the noise from the other bands, and we only plot the *grizy* points. These light curves are shown in Figure 2.6.

The transit in Figure 2.6 is not as visible as in the case of a Hot Jupiter, but can be discerned in the deep-drilling field where there are enough observations for the transit to be seen despite the noise. We provide the BLS periodograms for both a regular LSST field and a deep-drilling field in Figure 2.7, and indicate the true period. For the regular LSST field BLS does not recover the period and there are no strong discernable features in the periodogram that would correspond to the planet period. However, BLS does recover the input period for the deep-drilling field. This peak is higher than all but one BLS period from the 1000 permutated light curves, and a false positive rate of 0.1%.

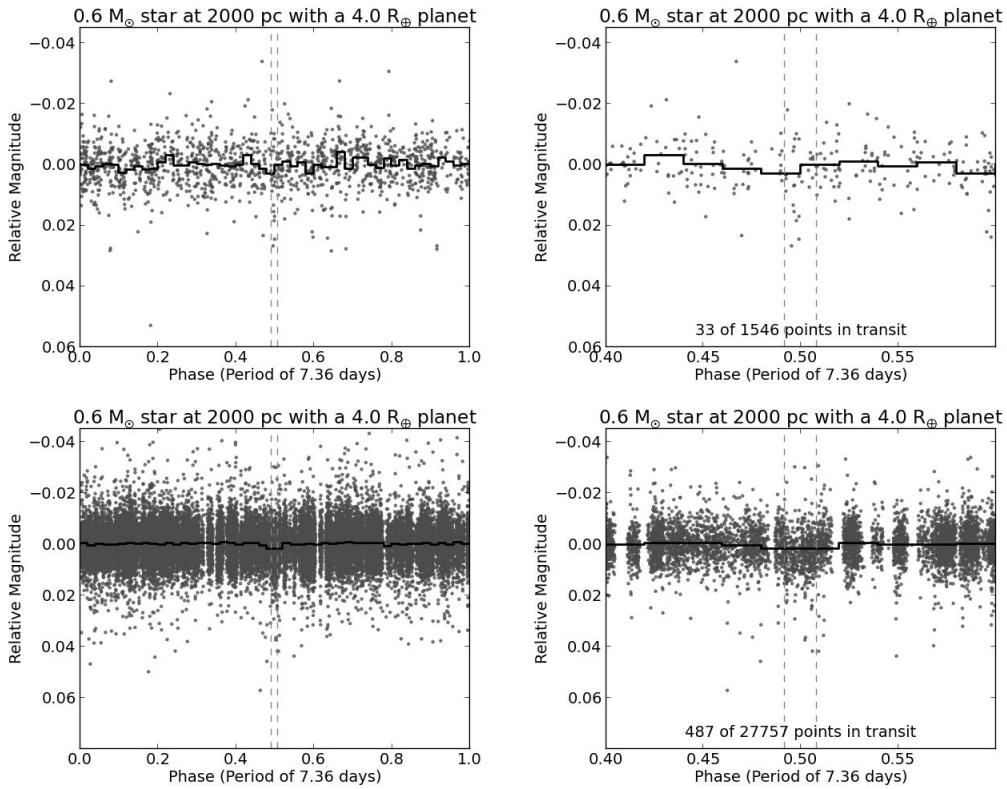

Figure 2.6: A 4.0 $R_\oplus$ planet in a 7.36 day period around a 0.6 $M_\odot$ star at 2000 pc. The top two plots show a regular LSST field and the bottom two plots show an LSST deep-drilling field. The plots on the left show the full phase of the planet, and the plots on the right show the transit in particular. Black lines are binned data of the light curve.

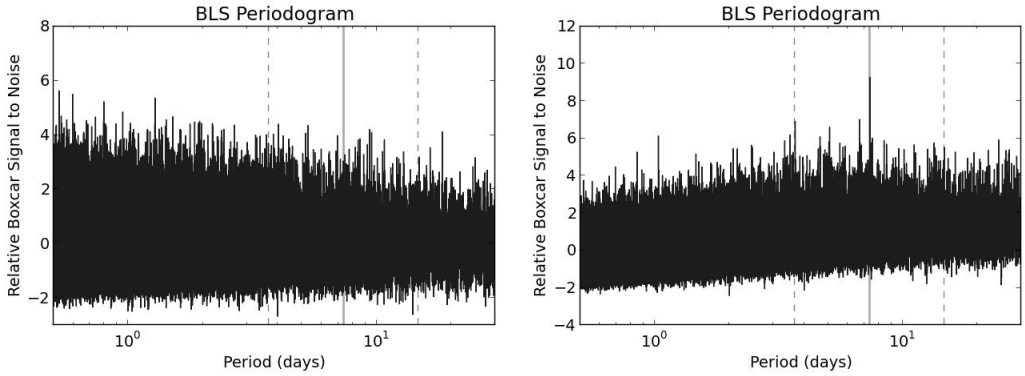

Figure 2.7: BLS periodogram for a Hot Neptune in a standard field on the left, and a deep drilling field on the right. The lighter grey line marks the actual period of 7.36 days, while the two dashed lines mark twice and half the true period.

2.4.3 A Super-Earth Around an M-dwarf in the Milky Way

For our final Milky Way case, we pick a 2.0 $R_\oplus$ planet in a 5.37 day period around a 0.3 $M_\odot$ star at 400 pc. The distance has again been adjusted so that the apparent magnitude of the red dwarf is approximately still within LSST's sensitivity range. As in the previous case, the *u*-band is omitted. In Figure 2.8 we show the light curves for the transit in both a standard field and a regular field. While in the standard field the transit is not visually detectable, in the deep-drilling field a small but slightly noticeable transit event is visible in the 5 remaining bands. With what appears to be a detectable signal in the deep-drilling field, we further explore a red dwarf host star with a Super-Earth in a larger period, moving the planet out to a period of 24.37 days. Our motivation in choosing this period is that for a red dwarf of this mass, a planet in a ~25 day period will be in the habitable zone, and this planet's system environment would be similar to that calculated for some planets in the Gliese 667C system [Gregory, 2012]. The light curve for this case is presented in Figure 2.9.

Figure 2.10 shows the BLS periodograms for both the 5.37 and 24.37 day periods, and at both cadences.

In both cases the input period corresponds to the highest peak in the periodogram[4]. In the 1000 permutations of the simulated habitable Super-Earth light curve in a deep-drilling field, we find that there are 390 higher peaks, corresponding to a false positive rate of 19.5%. Looking at these periodograms, however, that there is still potentially useful information present in the BLS results. While there appears to be no compelling features in the two lefthand figures that represent the two periods in regular fields, there are notable features in the deep-drilling fields on the right. For the 5.37 day period, the two highest peaks are the initial period and half the period. Similarly, for the 24.37 day period the

[4]We note that in this case, the BLS SN statistic identified the top peak as being one of the secondary peaks in the periodogram (Fig. 2.10) even though it is clear that the input period does coincide with the highest peak in the periodogram. We will examine the relevant differences between the various peak identification statistics and the BLS periodogram in a future paper.

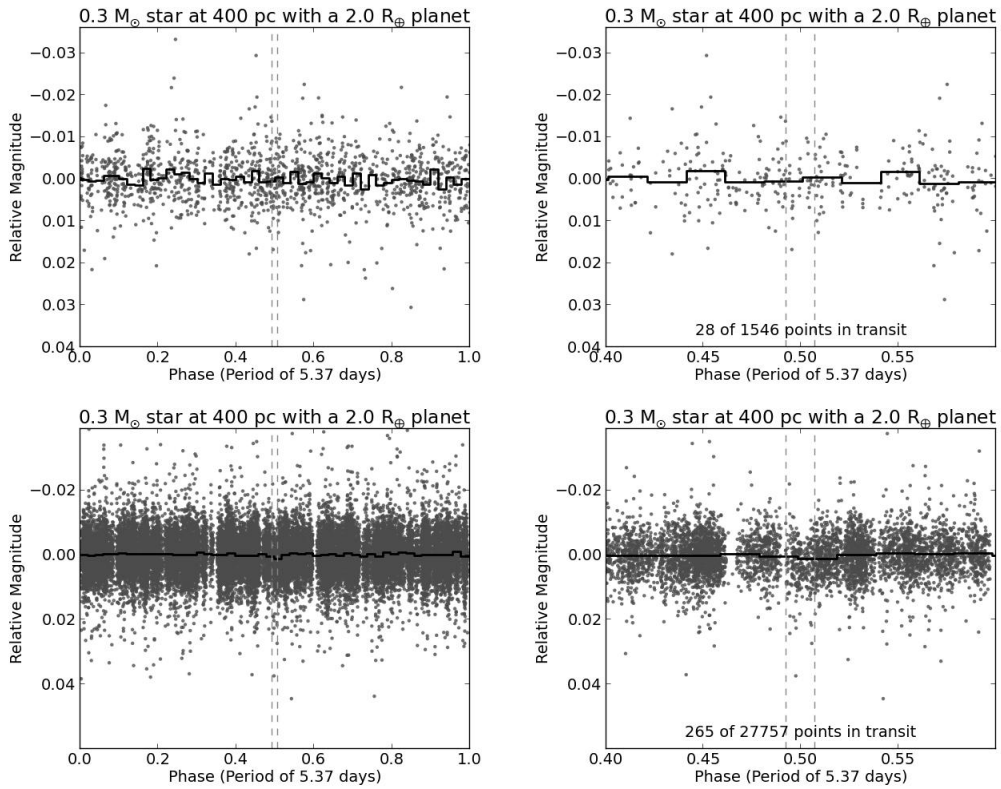

Figure 2.8: A 2.0 $R_\oplus$ planet in a 5.37 day period around a 0.3 $M_\odot$ star at 400 pc. The top two plots show a regular LSST field and the bottom two plots show an LSST deep-drilling field. The plots on the left show the full phase of the planet, and the plots on the right show the transit in particular. Black lines are binned data of the light curve.

initial period is one of the highest peaks, and there is a notable peak at half the period as well. While these period recoveries lack the rigor of being the top peak recovered by BLS, this does provide an indication that the transits could still be more reliably recovered in deep-drilling fields with the application of additional methods.

2.4.4 A Hot Jupiter in the Large Magellanic Cloud

LSST will also provide a unique opportunity to look for exoplanets outside of the Milky Way. Fields that have been proposed for LSST to observe at higher cadence include the Large Magellanic Cloud [Szkody et al., 2011]. Based on the faintness that LSST will be capable of observing, we demonstrate here that LSST should have the potential to be

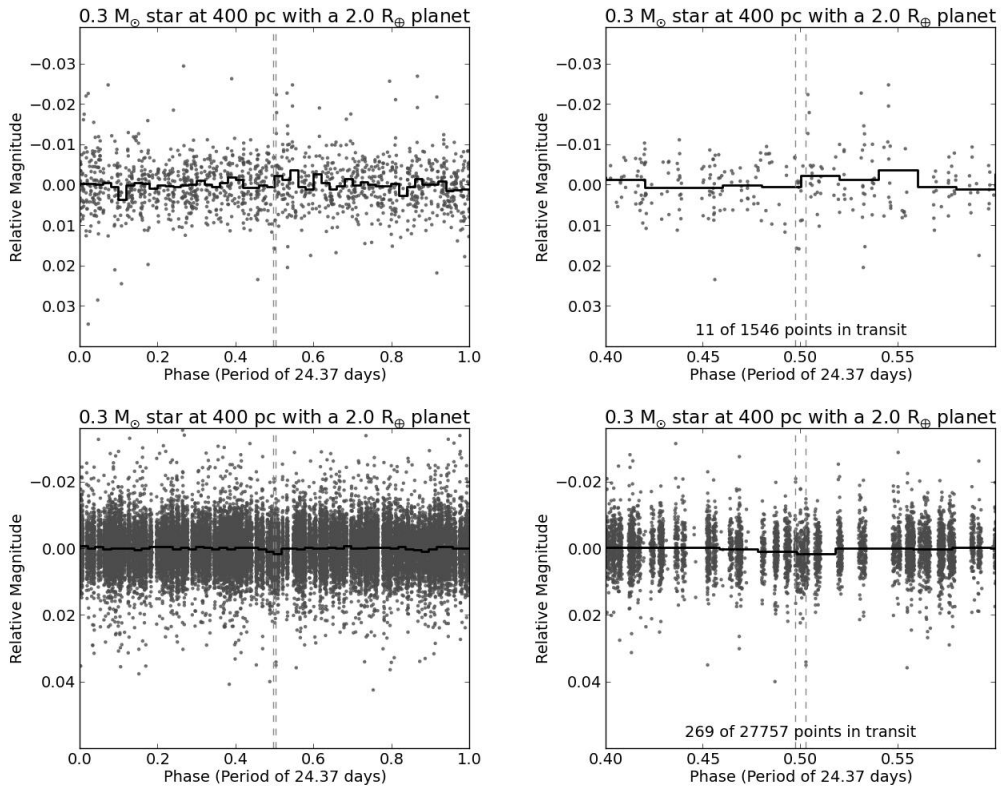

Figure 2.9: A 2.0 $R_\oplus$ planet in a 24.37 day period around a 0.3 $M_\odot$ star at 400 pc. The top two plots show a regular LSST field and the bottom two plots show an LSST deep-drilling field. The plots on the left show the full phase of the planet, and the plots on the right show the transit in particular. Black lines are binned data of the light curve.

sensitive to Hot Jupiters transiting stars located in the Large Magellanic Cloud. For this example, we look at a slightly larger Hot Jupiter than before, here a 13 $R_\oplus$ planet with a 4.4 day period orbiting a 1.0 $M_\odot$ star at 50,000 parsecs from Earth in Figure 2.11. As in the previous two examples, the *u*-band data has been omitted, but at this faintness the noise is sufficiently large that we also ignore the *gzy*-bands and only plot the *ri*-bands. While we are unable to detect the transit in the regular cadence light curve, we can detect it in the deep drilling light curve, as shown in Figure 2.12.

The entire set of deep-drilling fields has yet to be chosen, although the first four deep-drilling fields to be selected from LSST white papers have been intended for observations of distant galaxies [Gawiser et al., 2011, Ferguson, 2011]. However, other deep-drilling fields

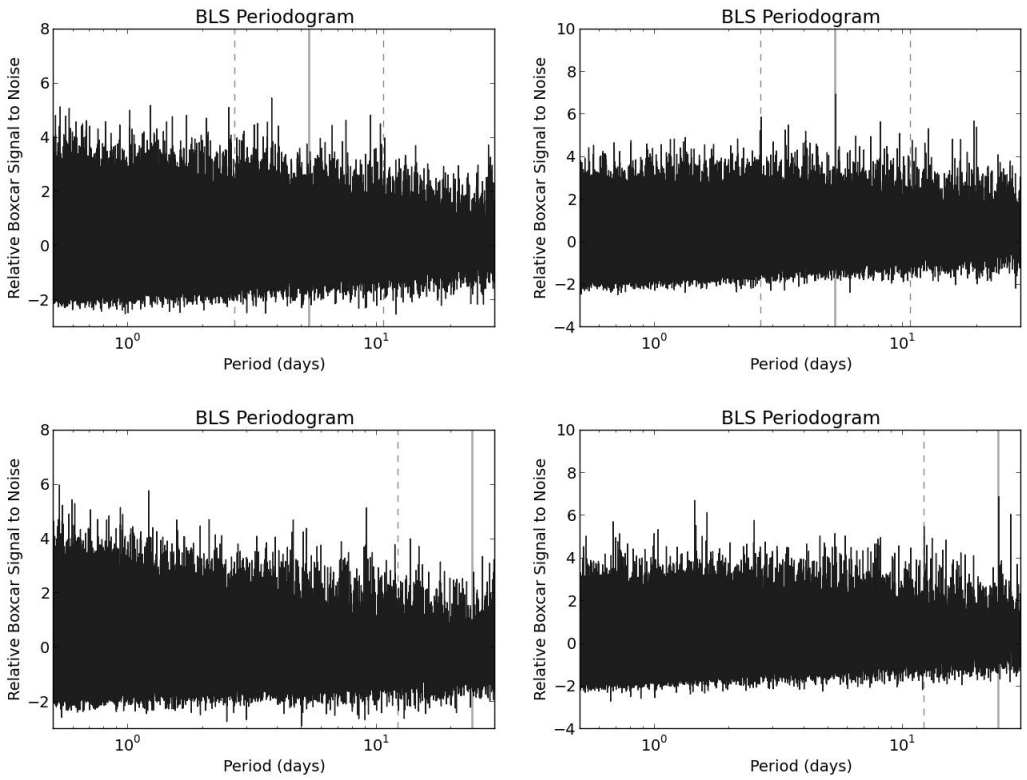

Figure 2.10: BLS periodograms for a Super-Earth in a regular field on the left and in a deep-drilling field on the right. The 5.37-day period is on the top and the 24.37-day period is the bottom. The lighter grey line marks the actual period for each scenario.

have also been suggested with the intention of focusing on the Large and Small Magellanic Clouds [Szkody et al., 2011]. These fields for the LMC and SMC would have an observing schedule which limits the filters used, and is proposed to only use *gr* band observations. This would be consistent with the limitations on photometric precision in the final light curves, which do not include several bands that we omitted due to large noise, and may result in light curves with much better photometric precision.

For a Hot Jupiter in the Large Magellanic Cloud, we provide the BLS periodograms for both a regular LSST field and a deep-drilling field in Figure 2.12. BLS does not recover the period in a regular field, nor do there seem to be any notable features in the periodogram, however we do recover the period in the deep-drilling field with a peak height that indicates a false positive rate of less than 0.1%. This result demonstrates that LSST could conceiv-

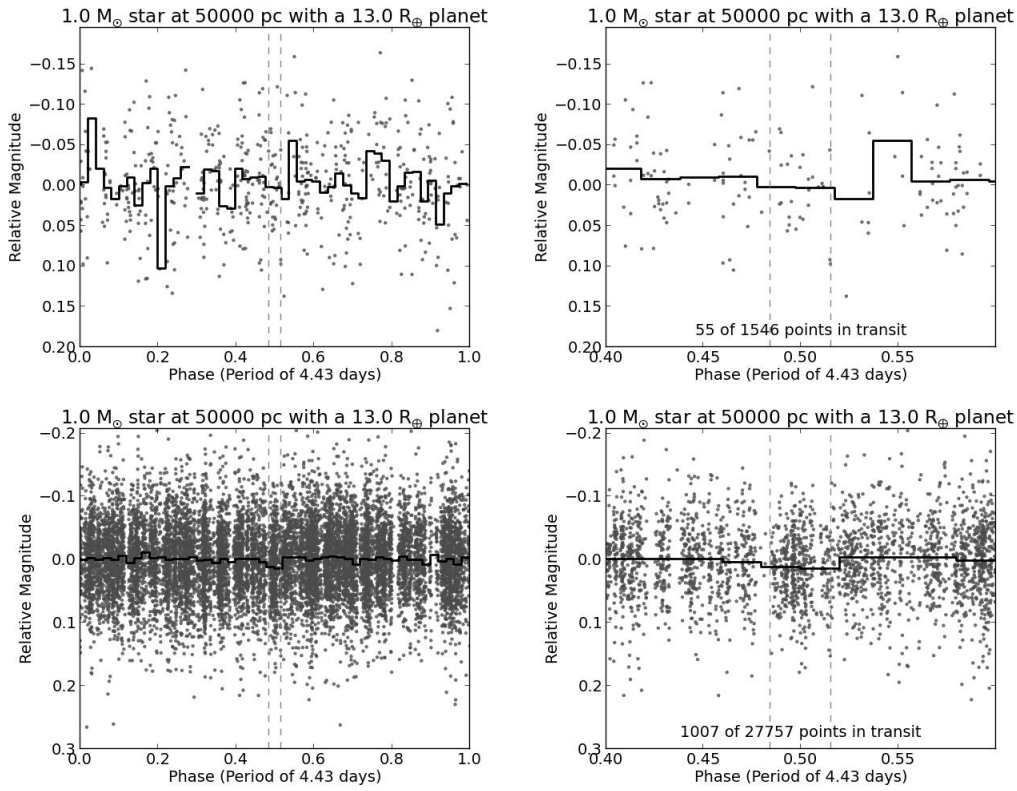

Figure 2.11: A 13.0 $R_\oplus$ planet in a 4.43 day period around a 1.0 $M_\odot$ star at 50,000 pc. The top two plots show a regular LSST field and the bottom two plots show an LSST deep-drilling field. The plots on the left show the full phase of the planet, and the plots on the right show the transit in particular. Black lines are binned data of the light curve.

ably detect a planet candidate outside the Milky Way. While it is worth noting that at the distance of the LMC it would be very difficult to confirm an exoplanetary candidate, this would certainly be a novel discovery and give some broader indication of planet formation processes in other environments.

2.5 Discussion

We have demonstrated that using existing algorithms, LSST will have the sensitivity to detect exoplanet transits for an array of exoplanets and host stars, and at varying distances. These results are summarized in Table 2.2. However, these capabilities will be greatly shaped by the selection process for the remaining deep drilling fields, as the selection of the

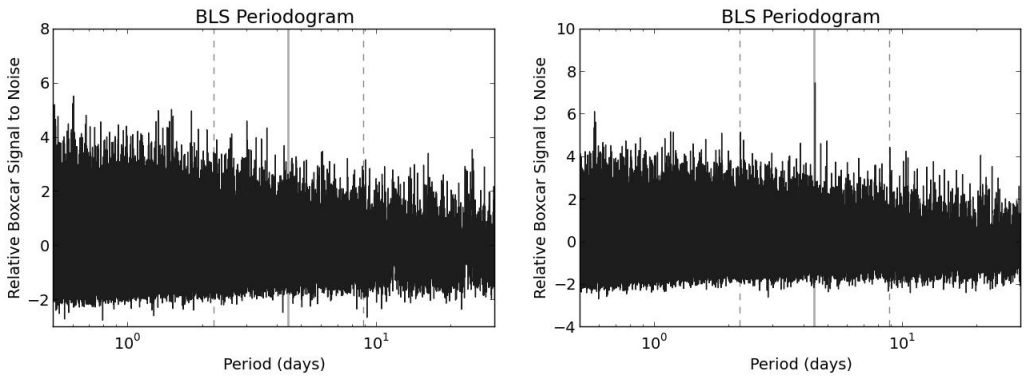

Figure 2.12: BLS periodogram for a Hot Jupiter orbiting a star in the LMC in a regular field on the left and a deep-drilling field on the right. The lighter grey line marks the actual period.

LMC or SMC for deep drilling will greatly increase the ability to get visually identifiable transits in those light curves.

Table 2.2: Scenario Detection

	Scenario			
Cadence	G-dwarf at 7000 pc Hot Jupiter	K-dwarf at 2000 pc Hot Neptune	M-dwarf at 400 pc Super-Earth (4.4 d)	M-dwarf at 400 pc Super-Earth (25.3 d
Standard	Yes	No	No	No
False Positive Probability	$< 10^{-3}$	–	–	–
Deep-Drilling	Yes	Yes	No	No
False Positive Probability	$< 10^{-3}$	10^{-3}	–	–

Table 2.2 lists whether or not the best period from VARTOOLS BLS algorithm is consistent with the input period for each of the scenarios described in §2.4 (and listed at the top of each column). The rows represent a standard cadence field and a deep-drilling field. Also included is the false positive rate when the period was correctly recovered by BLS.

For the case of Hot Jupiters, specifically, we can make a very preliminary approximation of the number of planets that could be recovered by LSST. Restricting our focus to the *g*-band, we find the number of stars between 16th and 22nd magnitude in a one-square-degree field at around 1800 stars, according to TRILEGAL 1.6[5]. There will, additionally, be around 30,000 square degrees observed, for a total of 54 million stars. Of these, we

[5] Available at http://stev.oapd.inaf.it/cgi-bin/trilegal

would expect around 0.5% of dwarf stars to host Hot Jupiters, and the likleihood of a Hot Jupiter to transit its host star is $\sim$ 10%. This would result in 27000 transiting Hot Jupiters, and recovering just 5% of these Hot Jupiters would result in over 1000 new Hot Jupiters discovered. While these are very approximate numbers, they do indicate the large potential that LSST possesses.

Through our work thus far, we have treated host stars as having no intrinsic variability, however a more sophisticated model of stellar light curves will be needed going forward. Stellar variation from rotation, pulsation, flaring, and eclipses will need to be included to better synthesize the kinds of light curves that will likely be observed by LSST, and these variations will complicate recovering the signal caused by a transiting planet and be a source of false negatives. Additionally, we will need to incorporate these light curve features into the light curves of stars in the field without planets in order to properly model false positives.

A greater difficulty will be developing methods that can more rigorously detect and recover these transits, and determine the difference between likely exoplanetary candidates and false positives, astrophysical or otherwise. Due to the low cadence of the data, standard methods used for finding periodic signals in light curves are potentially not sufficient to find transits in all circumstances, and newer algorithms will need to be implemented with these constraints in mind. The efficiency of these algorithms will be an important question, both to quantify the usefulness that LSST will have in detecting extrasolar planets, and in determining the best way to structure a search for transiting planets in the LSST data once observations begin in 2020.

There will also be a need to expand this study from the discrete example cases we have set forth here, and apply similar analysis to a more diverse set of systems. That will include expanding our synthetic light curves to represent a distribution of stellar masses and exoplanet radii and periods, as well as factoring in the likelihood that a planet will transit. This will allow our results to look not just at the likelihood for a given planet to be

detected, but also to determine how many exoplanets could reasonably be detected among stellar populations such as field stars and star clusters, and will also characterize the variety of planet types that will be detectable using the LSST data.

While we have discussed potential opportunities for exoplanet detection specifically from LSST, these same methods can be applied to other surveys that have a large number of observations over relatively long baselines, but with similarly sparsely-sampled data that will make extracting exoplanetary transits difficult without implementing more novel algorithms. Surveys that would warrant such consideration may include Pan-STARRS [Kaiser, 2004], the Catalina Real-Time Sky Survey [Djorgovski et al., 2011], the Palomar Transient Factory [Rau et al., 2009], and GAIA [Jordi and Carrasco, 2006]. Once our procedures for stellar and exoplanetary population simulation, light curve creation, and application of detection algorithms have been developed for LSST, we will then be able to apply these same techniques to these other surveys to greatly increase the data available to search for exoplanets in varied situations.

Chapter 3

Metrics for Optimization of Large Synoptic Survey Telescope Observations of Stellar Variables and Transients

3.1 Introduction

The Large Synoptic Survey Telescope (LSST) is expected to provide time-domain photometric monitoring of billions of stars, over a period of ~ 10 yr and with unprecedented breadth and depth. This will in turn permit an unprecedented ability to identify and characterize a variety of stellar variability and transient phenomena, from non-recurring novae and supernovae to periodic variables such as eclipsing binaries. It will also more generally probe a very broad array of stellar types and populations.

To more fully capitalize on the potential of the LSST to contribute to a broad range of astronomical interests, the LSST team has publicly released data and tools so that the wider community can contribute to discussions of how to optimize the potential of LSST for any science outside of LSST's stated goals. The first of these tools is the LSST Operations Simulator (OpSim) which models the hardware and software performance, site conditions, and cadence as a function of sky position [Delgado et al., 2014]. The other tool that has been provided by the LSST team is the Metric Analysis Framework (MAF)[1][Connolly et al., 2014]. The objective of the MAF is to allow the comparison of different simulated LSST data sets by reducing the properties for each field or position in the sky to a single number which can then be mapped.

In this paper[2], we introduce three MAF metrics[3] for use with LSST. These metrics are intended to address three questions:

[1] Source code at https://github.com/lsst/sims_maf. Full installation guide at https://confluence.lsstcorp.org/display/SIM/MAF+documentation

[2] This chapter was originally published as [Lund et al., 2016b]

[3] Available at https://github.com/LSST-nonproject/sims_maf_contrib

- How well can a given observing plan detect transient photometric phenomena with constant equatorial coordinates?

- With what level of completeness and precision can a given observing plan detect periodic photometric phenomena with constant equatorial coordinates?

- How many main sequence stars in the Milky Way will be detected in each specific LSST observing field, and what will be the range of their masses?

In §3.2 we summarize the basic LSST specifications and simulation tools that we use. In §3.3 we present the three metrics and describe how they can be used to characterize a given observing plan (cadence, coverage, depth) toward optimization for various science applications. Then in §3.4 we review the combined usefulness of the metrics presented here and describe how the development of further metrics can enhance the science yield of the LSST.

3.2 Basic LSST Specifications and Tools Used

The Large Synoptic Survey Telescope will collect ten years of photometric observations in six bands over more than half the sky, scheduled to begin observations in 2020. LSST has the capability to contribute greatly to many areas of science, but there are four goals explicitly stated: Taking an Inventory of the Solar System, Mapping the Milky Way, Exploring the Transient Optical Sky, and Probing Dark Energy and Dark Matter [LSST Science Collaboration et al., 2009]. The LSST system design has been dictated by these goals, and hardware specifications have already been established. The need to be able to survey the entire visible sky with relatively short (less than thirty seconds) exposures has resulted in a telescope with a relatively large field of view of 9.6 square degrees.

The observing schedule that will be used has not yet been finalized, however, and is in many ways still open to suggestions for how to best optimize the survey. At this point, LSST observations can be quantized as individual 'visits', consisting of two consecutive

15-second exposures in the same band and location, but this could be changed if a compelling science case was presented. The survey can broadly be divided into a few distinct subsurveys, with the primary 'deep-fast-wide' survey being what we refer to in this paper as regular cadence fields. These fields will have ~1000 observations over the 10 years of LSST's mission in 6 bands. The other large component we discuss is the 'narrow and deep' survey, or what we refer to as the deep-drilling fields. These will be ~10 fields that will each be observed ~10,000 times in 6 bands, primarily consisting of 40 one-hour block observations for each field. The LSST survey also includes subsurveys that will focus on the galactic plane, polar cap, and the North Ecliptic Spur.

In this paper we utilize two tools made available by the LSST team: The Operations Simulator (OpSim) and the Metrics Analysis Framework (MAF). For our purposes, OpSim produces a list of observation times with field ID, band, and limiting magnitude. For this paper we use two OpSim data sets; opsimblitz2_1060 (also referred to as ob2_1060) consists of several subsurveys and has been introduced as a potential reference run, and ops2_1078 only consists of the 'deep-fast-wide' regular cadence fields, but with different cadence parameters[4]. The number of observations for both fields are shown in Figure 3.1.

Figure 3.1 is the output produced by the MAF for a simple metric that calculates the number of observations per field. The LSST team has already provided several sample metrics as part of the initial MAF package. However, third parties can also create independent metrics. This capability allows for the community to provide new metrics to address properties of the survey tied to the observability of a myriad of scientifically interesting objects. In what follows, we use the LSST MAF to demonstrate three new metrics that permit quantitative analysis of the performance of different OpSim designs for the survey cadence and coverage, specifically focused on optimizing LSST's coverage of different stellar populations, variables, and transients.

[4]Available at https://confluence.lsstcorp.org/display/SIM/

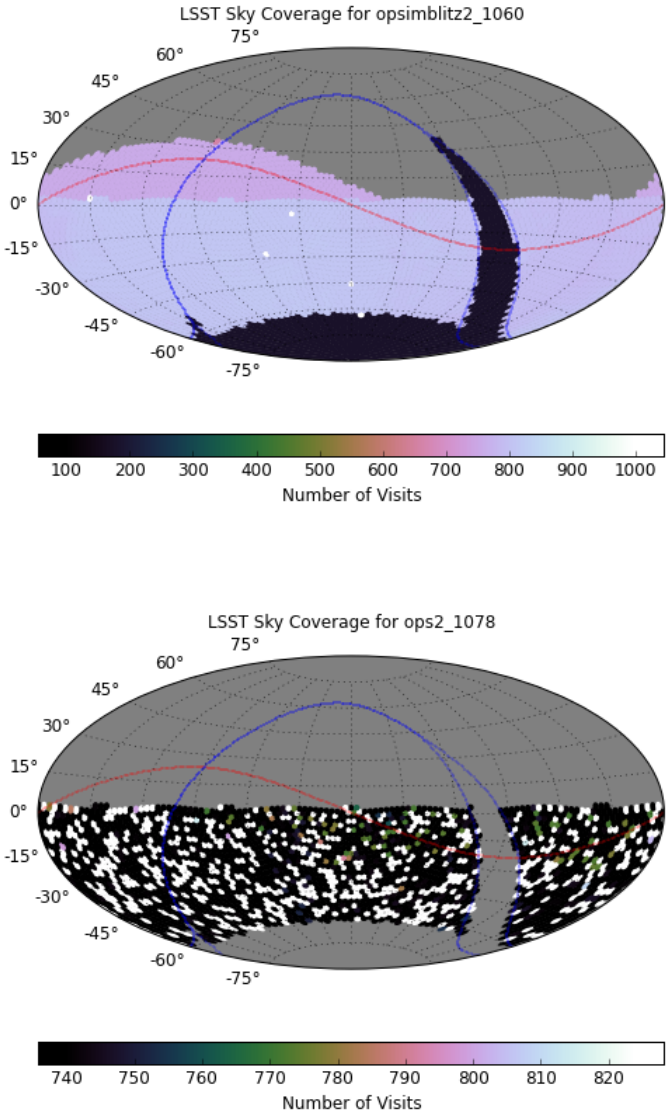

Figure 3.1: An example of a metric as displayed by the Metric Analysis Framework, the total number of observations per field. Opsimblitz2_1060 on the top, and ops2_1078 on the bottom. Opsimblitz2_1060 shows the primary survey, in addition to the deep-drilling fields and additional subsurveys. Ops2_1078 only features the primary survey.

3.3 New Metrics for Stellar Populations, Variables, and Transients

3.3.1 Observation Triplets: Nonrecurring stellar transients

A basic concern for LSST centers on its ability to detect transient events. However, many of the metrics that have already been produced for the Metric Analysis Framework

focus on coadded images rather than considering how to detect time-dependent photometric variability. This broad category of potential transient events ranges from asteroid collisions to stellar novae and stellar mergers. As all these events require observations during a fixed time interval, an important question for LSST will be how survey cadences will impact LSST's ability to detect events that may take place on timescales from hours to days.

For LSST to suitably detect transient events, we need to define the time resolution necessary to detect a given event. Here, we consider an event that consists of a relatively constant baseline flux, followed by a change to a different flux which remains relatively constant for a period of time. An example of such a transient event is a nova that brightens after a long time of constant brightness, without any simple periodicity. In order to consider this a detection, we require one observation prior to the event, a second observation at the start of the event showing the change in brightness, and a third observation that confirms that the change in brightness wasn't due to noise in a single data point. From these three data points, the magnitude change can be measured and the time of the event can be constrained between the first and second observations, allowing for basic detection and characterization. We can then quantify the ability of a given LSST field to detect non-periodic events by counting the number of these "observation triplets" that occur.

To provide an example, we look at the case of detecting a nova. In Figure 3.2, we show a simplified light curve of the nova SS Cyg based on AAVSO observations [Kafka, 2015]. The vertical line, T2, marks an ideal time for the middle observation of the observation triplet, shortly after the brightness has increased. We then require that the interval between the first and second, and second and third, observations to fall somewhere between Δ_{min} and Δ_{max}, and these regions are shaded in red in the figure. Two potential observations are depicted by the vertical lines T1 and T3, with Δ_{12} and Δ_{23} being the time differences between the first and second, and second and third observations, respectively. We can also put limits on the ratio of these two time intervals.

This metric is written so that Δ_{min}, Δ_{max}, and the maximum ratio of Δ_{12} and Δ_{23} are

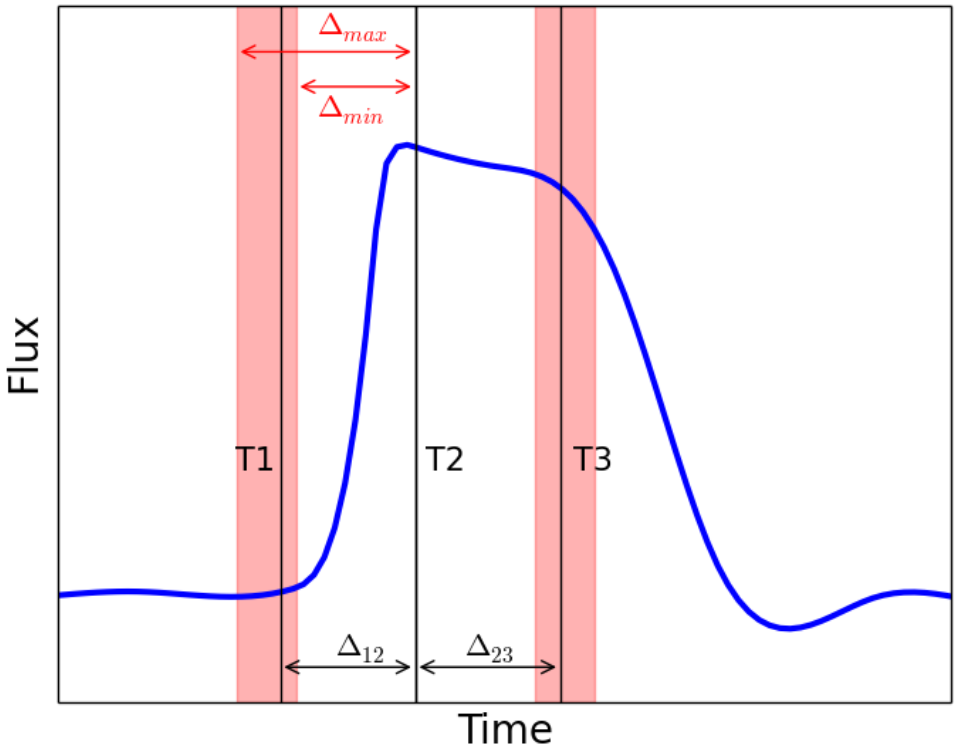

Figure 3.2: An illustration of a basic light curve of a non-periodic brightening. T1, T2, T3 represent a set of observations that would satisfy our criteria to detect this event. Using T2 as our reference point, T1 and T3 could fall anywhere within the regions shaded in red that are bounded by a minimum and maxmium time interval.

parameters of the metric. For our example, we set the minimum time interval between observations at 9 days, the maximum at 11 days, and the two time intervals must be within 10% of one another. In essence, this asks how many opportunities LSST will have to observe non-repeating brightening or dimming events that occur on various timescales. Note that the nova example is intended to be illustrative only; the point here is to have a metric that captures the sensitivity to transient behavior on various timescales, which the user can modify as needed for the timescale(s) and behavior(s) of interest. We also require that the observations of a given triplet all be made in the same band.

Figure 3.3 shows the results we get for two different OpSim runs. In the first OpSim

run the deep drilling fields show up strongly with the number of observation triplets in these fields exceeding 100,000. Additionally, the subsurvey that covers the North Ecliptic Spur, a 4,000 degree area along the ecliptic in the northern hemisphere that has larger airmasses than the main survey observations, also has more observation triplets than the primary deep-wide-fast survey, as do portions of the galactic plane. In the second OpSim, the different cadence results in the number of observation triplets in the main survey being slightly higher, although gaining less than an order of magnitude per field. For the purposes of detecting the novae in this example case, the cadences used in the deep drilling fields, North Ecliptic Spur, and portions of the galactic plane will be far more effective than the cadence of the main survey. The majority of the sky will be surveyed at the regular 'deep-wide-fast' cadence, however, and in the main survey the cadence that is used in OpSim results ops2_1078 is slightly more beneficial for detecting novae; in ops2_1078 the mean number of observation triplets per field is 5.3, compared to an average of 3.0 observation triplets for the 'deep-wide-fast' cadence fields in opsimblitz2_1060.

While we have not explored it in this paper, the Metric Analysis Framework does also allow for analyzing the survey using positions in the sky rather than individual fields. This method would allow for areas of the sky where fields overlap to be counted multiple times, and may be particularly useful at timescales that are on the order of an hour or less. Individual fields will generally not be observed closer than one hour apart, but overlap regions may be observed more frequently.

3.3.2 Periodogram Purity Function: Detectability of Periodic Variability

A basic question for LSST is how robustly it can detect periodic behavior. In particular, choosing a cadence for this sort of detection requires some special consideration as the ideal cadence will be one that minimizes the effect of aliasing. To address this, we directly investigate aliasing as caused by the LSST cadence.

LSST will observe numerous kinds of transient periodic objects, such as variable stars,

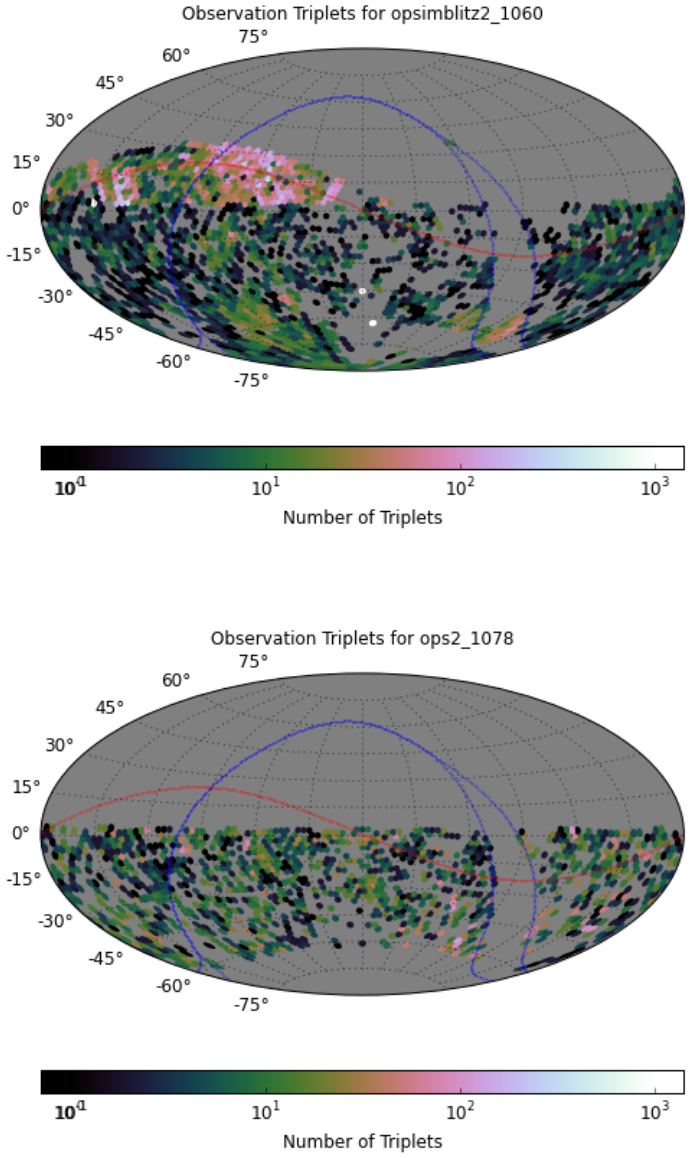

Figure 3.3: The number of observation triplets on an example time scale, with opsimblitz2_1060 on the top, and ops2_1078 on the bottom. The observation triplets are more numerous in fields in the North Elliptic Spur and deep-drilling fields shown in opsimblitz2_1060.

eclipsing binaries, and transiting exoplanets. While each of these kinds of events will have their own unique characteristics, we opt to establish a metric that examines the completeness of phase coverage for a range of periodic timescales, as a way of quantifying

LSST's relative likelihood to permit detection of any given periodic variable. Specifically, the metric calculates the Fourier power spectral window function of each field [Roberts et al., 1987]. The spectral window function provides a measure of how the finite and discrete cadence and sampling of a real light curve result in the Fourier periodogram power getting distributed across multiple alias and beat frequencies. For any real light curve, the resulting Fourier periodogram is a combination of the idealized, "pure" periodogram (i.e., a delta function at one frequency in the case of simple sinusoidal variations) and the spectral window function which "leaks" that pure power into many other periodogram peaks, decreasing the power of the true frequencies and decreasing the "purity" of the overall power spectrum.

To create our periodogram purity function metric, we calculate the Fourier spectral window function , F_w, for a given light curve sampling pattern, and we then define the periodogram purity function as $1 - F_w$. In the ideal case with no gaps in observation, the spectral window function can be represented as a normalized Kronecker delta centered at 0 and having a height of 1, meaning that all the Fourier power is contained precisely at the correct period and no power is leaked to other frequencies. The periodogram purity function would in this ideal case equal 1 at all nonzero frequencies. As phase coverage is reduced from the ideal case to the real case of finite and discrete data, signal strength will be transferred to other frequencies, creating greater structure in the periodogram. In order to quantify the amount of signal being lost for a given set of observations as a single number for the LSST MAF, we look for the smallest value of the periodogram purity function away from the true frequency. Under that convention, this means that two sets of observations could be compared where the smaller the minimum value of the periodogram purity function is, the more power that has been lost from the actual period. That is, the less "pure" is the periodogram at frequencies away from the true frequency.

To demonstrate the value of the periodogram purity function when reduced to a single value, we create two sample light curves. In both cases, we begin with a sine curve with

a period of 4.5 days, and shown in Figure 3.4. To choose our observing times, we set 1000 evenly spaced observations over the course of ten years. We then add a random number to each observation time in order to have some variation in observing time intervals; in one case we add random values between 0 and 1 days (referred to hereafter as our wide distribution), and in the other case we add random values between 0 and 0.001 days (referred to hereafter as our narrow distribution). In the wide distribution, the time between observations ranges from less than 3 to over 4.5 days. In the narrow distribution, the time between observations ranges only from 3.645 to 3.665.

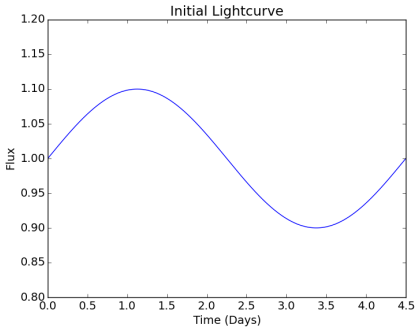

Figure 3.4: We use a standard sine function of the equation $y = 0.1 * sin(x * 2 * \pi/4.5) + 1$ for our light curves that we analyze in Figure 3.5 and 3.6.

To compare these two light curves, we calculate the Lomb-Scargle periodogram[Lomb, 1976, Scargle, 1982] and the corresponding periodogram purity function, and these are displayed in Figure 3.5. The periodogram for the narrow distribution shows numerous peaks of comparable power, with 13 peaks within 1% of the strength of the top peak. In contrast, for the wide distribution the highest peak correctly corresponds to the period of the sine curve we added to the light curve. Qualitatively speaking in this case, the wide distribution appears to be much better at determining the period of the light curve. When we compare the periodogram purity function, we see that for the narrow distribution that poorly recovered the period, the lowest peak at nonzero frequency shift is less than 0.02. The wide distribution that does recover the period correctly has its lowest peak at just over 0.91, demonstrating that the light curve that was able to more definitively identify

the period and have greater phase coverage was also the one with the periodogram purity function value much closer to 1. These periodogram purity function values are agnostic with respect to the nature of the periodic signal we are interested in. An important note is that the spectral window function, and in turn the periodogram purity function, represents how power is distributed to other frequencies from the true period. Changes to the periodic function we are interested in, such as variations in the amplitude or different functional forms, will impact the power that is present; however, the redistribution of this power can be represented by the same term for a fixed set of observation times.

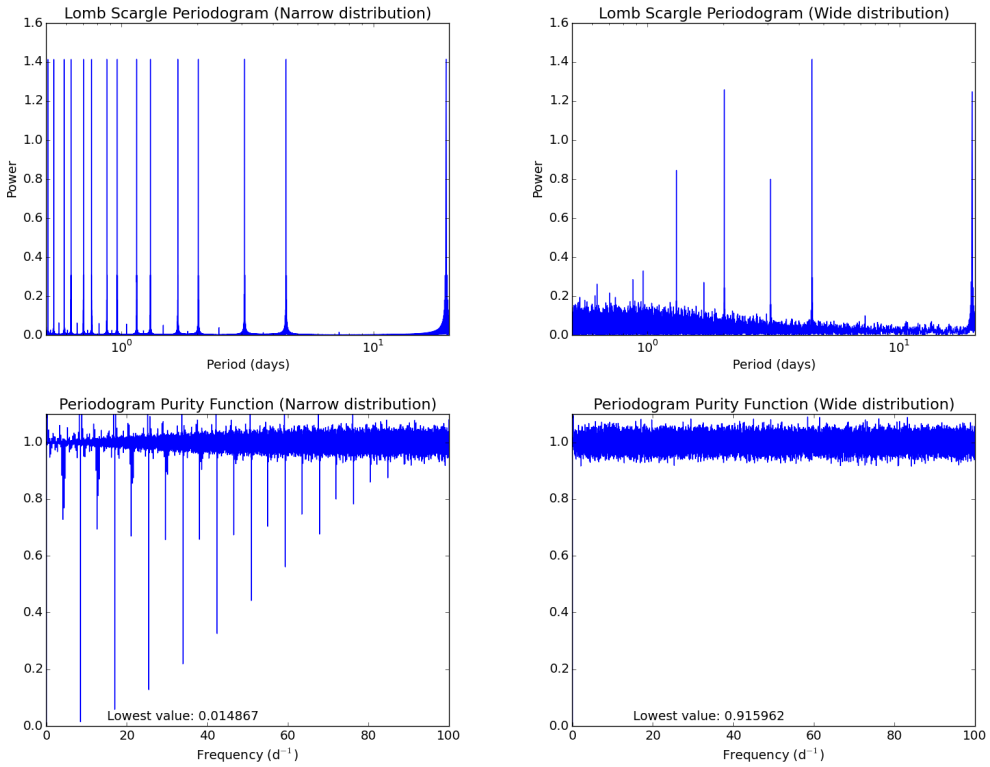

Figure 3.5: The top row shows the Lomb-Scargle periodograms for the two sample light curves described in §3.3.2. In the periodogram, the peaks represent potential periods of the light curve, including aliases. The bottom row shows the corresponding periodogram purity functions. In an ideally-sampled light curve, the periodogram purity function would be 0 at 0 frequency and 1 at all other values.

We can build upon the sample cadences in Figure 3.5 by looking at two actual cadences from the OpSim results. We choose a regular field consisting of on the left in Figure 3.6,

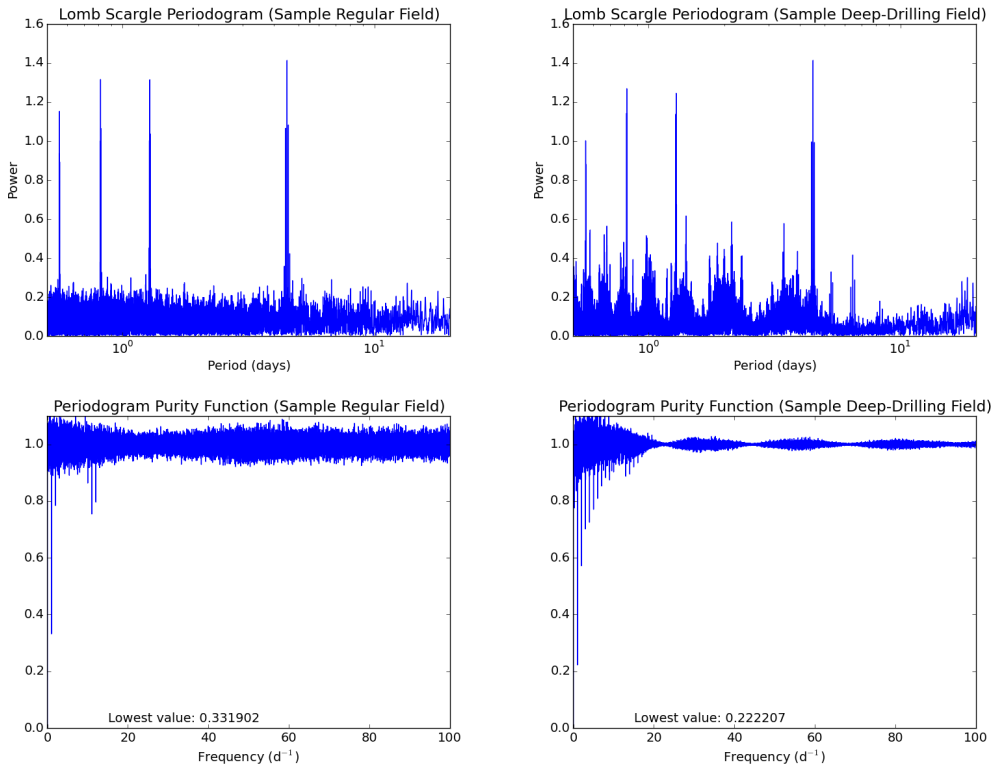

Figure 3.6: The top row shows the Lomb-Scargle periodograms for two fields from the OpSim results; on the left is a regular field and on the right is a deep-drilling field. The bottom row shows the corresponding periodogram purity functions.

and a deep-drilling field on the right. It is worthwhile to note that for the metric we take the lowest value at a non-zero frequency for the value of the periodogram purity function, and there is very little difference in this value between a regular field and a deep-drilling field. However, unlike our initial test cases, there is interesting non-uniform structure found in the periodogram purity functions, with the deep-drilling fields having more scatter at lower frequency shifts, indicating that there is more loss at small changes in frequency, and we can expect to see broader peaks in the periodogram due to this.

Using the Metric Analysis Framework, we then look at the minimum value (at a non-zero frequency shift) of the periodogram purity function for every field in two observing simulations in Figure 3.7. A general characteristic that appears is that in the overall survey, each subsurvey represents a different regime of phase coverage. These results are depen-

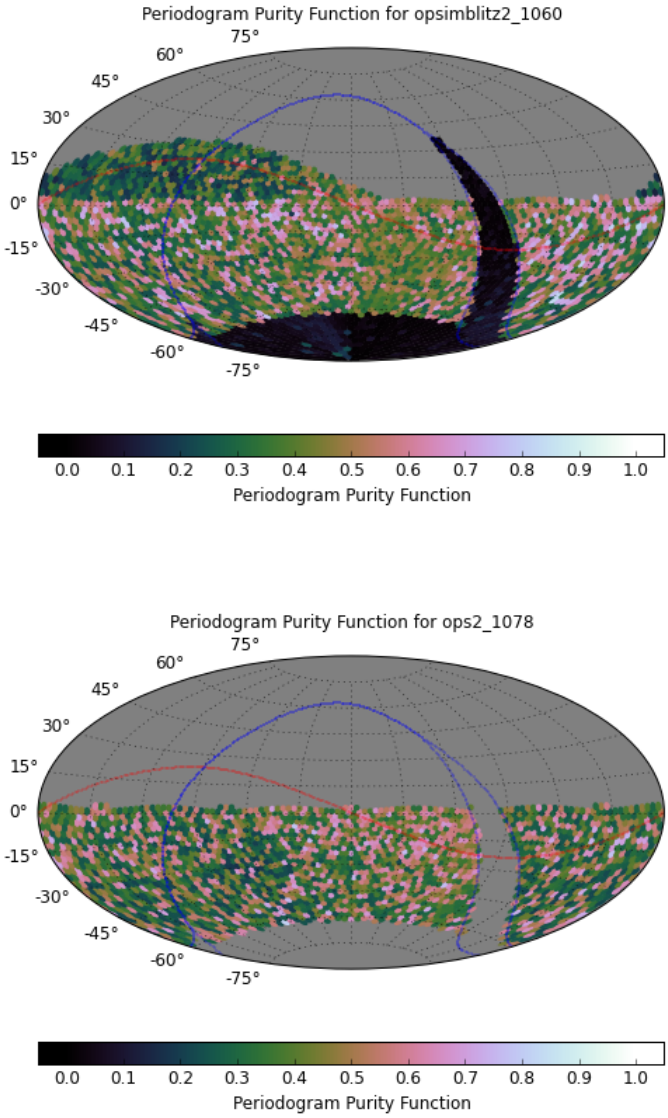

Figure 3.7: Periodogram purity function minimum values (at non-zero frequency shift) for each field in two simulated observing runs, opsimblitz2_1060 on the top, and ops2_1078 on the bottom. An ideal time sampling would have a periodogram purity function value of 1; this function is only dependent on the times of observation.

dent only on the observing schedule and are not dependent on any input signal. For the first OpSim result, we see that the main survey has the best phase coverage, and the other sub-surveys will have worse phase coverage. In the second OpSim result, the adapted cadence has approximately the same phase coverage for the main survey (a median periodogram

purity function value of 0.45 for the main survey in opsimblitz2_1060 and a median of 0.41 in ops2_1078).

For any observations that will require high quality phase coverage, we see that the south celestial pole and galactic plane do not provide any significant phase coverage. Additionally, the North Ecliptic Spur and the deep-drilling fields will have phase coverage equal to the main survey fields. The current metric we provide for use with the Metric Analysis Framework provides some insight, however there may be further benefit in an increased study of the structures shown in the periodogram in Figure 3.6 to develop a more sophisticated understanding of what phases LSST will be able to best recover, particularly in the deep-drilling fields.

3.3.3 Field Star Counts

There are two principal components that will need to be considered to understand the total yield of any transient object. The first component is the cadence, which is measured in the metrics presented in Sections 3.3.1 and 3.3.2. The other important factor is the number of potential sources in a given field, which for stellar sources would represent the number of stars that would be observed with sufficient signal to noise to be useful, while also being faint enough that the LSST detector will not saturate. In order to address this concern, we develop a metric to determine the number of stars in a given field that satisfy some set of observation parameters.

LSST will be able to observe a very wide range of stars with relatively low noise, including stars as small as red dwarfs, and distances ranging from within ~100 parsecs out to over 10,000 parsecs, including the LMC. This will be, in part, due to the multiple bands that LSST will observe in. We show this wide range in Figure 3.8, where we examine how many bands main sequence stars would be observable in. For our bright limit we use the apparent magnitude of 16 as a saturation point. For our faint limit, we required the photometric noise of a single visit to be less than 0.03 mag, a value that we used as a

cutoff for the maximum noise while still being able to detect a transiting planet in Lund et al. [2015], or an approximately 1% drop in brightness during transit. We convert from stellar masses to spectral types by interpolating the relations from Allen and Cox [2000]. Based on spectral type, we obtain absolute magnitudes in the *ugriz* bands from Covey et al. [2007]. While there is still much uncertainty about the specific parameters of the *y*-band filter that LSST will eventually use, for this work we use a *y*-band defined by Hodgkin et al. [2009]. This allows us to convert a set of stellar masses into a set of absolute magnitudes. The present version of the Field Star Counts metric does not employ any model to account for dust.

There are a few notable features of Figure 3.8. There are two excluded regions of magnitude/distance where LSST would not be sensitive at the specified thresholds. The empty upper left corner shows that all stars more massive than $\sim 1.5 M_{\odot}$ and closer than 1 kpc are saturated in all LSST bands. The lower right corner shows that all stars less massive than $\sim 0.7 M_{\odot}$ and more distant than 10 kpc will not be detected with the specified photometric precision for exoplanet transit detection in any LSST band. Running diagonally through the Figure, there is a very large range in stellar masses and distances where stars will be measured with sufficient sensitivity to detect small variations such as transiting planets. Stars with masses below 0.5 $M_{\odot}$ are observable within ~4000 parsecs in several bands, and as close as ~100 parsecs. Solar-mass stars are observable in at least 5 bands from ~2000 to ~20,000 parsecs, a distance range that includes the galactic bulge. Stars slightly larger than one solar mass that are in the Magellanic Clouds, at ~50,000 parsecs, would be observed within the specified constraints. Co-added images will allow the bottom right of this figure to be observed within this noise threshold, but without the time resolution that would be critical for short timescale variations in brightness.

As the range of observable stellar masses is extremely broad, there is a great benefit in determining the number of stars that satisfy these parameters for a given type of event. LSST covers a large region of sky, and so the method we use is designed to not be computa-

tionally intensive, or to rely on a more thorough and complex method such as the simulation process that is conducted by TRILEGAL [Girardi et al., 2005]. Instead, we only calculate the number of stars within a volume based on stellar density, rather than attempt to create a simulated population of stars.

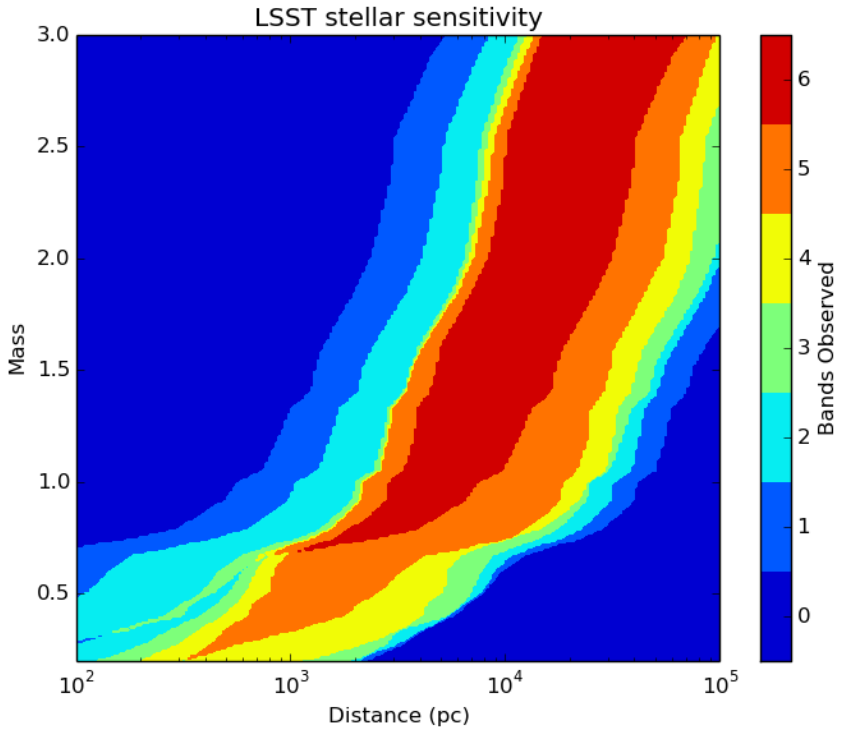

Figure 3.8: Number of bands in which a main sequence star will be observable between a saturation-limited upper brightness and a noise-limited lower brightness for a single visit.

The basic version of this metric calculates the number of stars within a given field and within a particular distance range. We accomplish this by treating each field as a truncated cone spanning two distances. We then break this cone into sections of equal volume and assign a stellar density for each volume as a function of the galactic coordinates of the center of that section. This begins with the analytical equation for the density of stars in the Galaxy, as a combination of the thin disk, thick disk, and halo, as presented in Jurić et al. [2008]. As Jurić et al. [2008] had focused only on the local portion of the Milky Way, that equation does not include a contribution from the galactic bulge, and so we include a term

for the galactic bulge from Jackson et al. [2002]. Without applying any brightness cut, we determine a stellar count in each field (integrated from 100 to 1000 pc as a demonstration) as shown in Figure 3.9.

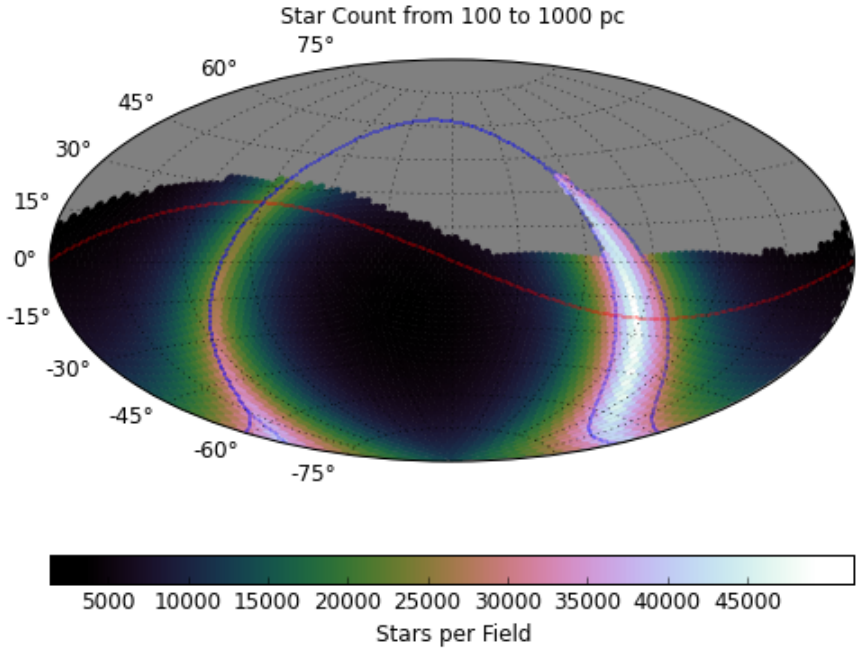

Figure 3.9: Stellar count per field between 100 and 1000 parsecs

To look at specific stellar masses, we define a mass range of interest. This set of stellar masses is then converted into a set of absolute magnitudes. Combining those values, the noise model, and the apparent magnitude limit, we can determine a distance range over which these stars would be within LSST's sensitivity for a set of stellar mass bins. We then calculate the total star count in the distance range that corresponds to each stellar mass bin, and then select only the stars with masses in that stellar mass bin by applying a standard Salpeter IMF. We demonstrate this technique with two example mass ranges, with stars between 0.4 and 0.5 solar masses in the *g* band in Figure 3.10, and stars between 0.9 and 1.0 solar masses in the *u* band in Figure 3.11. The Field Star Count metric can be used in conjunction with metrics that are developed for stellar observations, as a field that performs

well in a given metric may not be useful if there are many stars in that particular field.

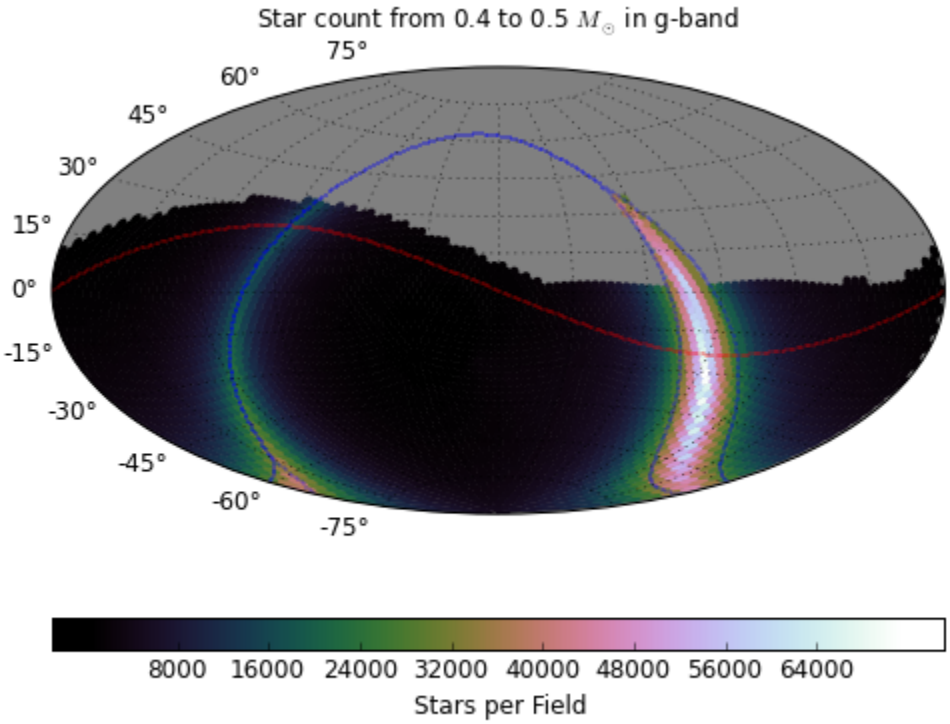

Figure 3.10: Star count per field for the 0.4-0.5 solar mass range and noise less than 0.03 mag. Observations in the *g*-band.

3.4 Summary

In this paper we have introduced three new metrics to help compare LSST observing strategies in advance of the start of LSST's mission in 2020. The Observation Triplets provide a tool to study any transient event that could be sufficiently discovered or characterized by an observation at a baseline brightness and two observations detecting and confirming the event. The tool can be adapted for a range of timescales, however we provide the specific example of a ten-day timescale that would be suitable for stellar novae. For this example, the deep-drilling fields and the North Ecliptic Spur would be the cadences that

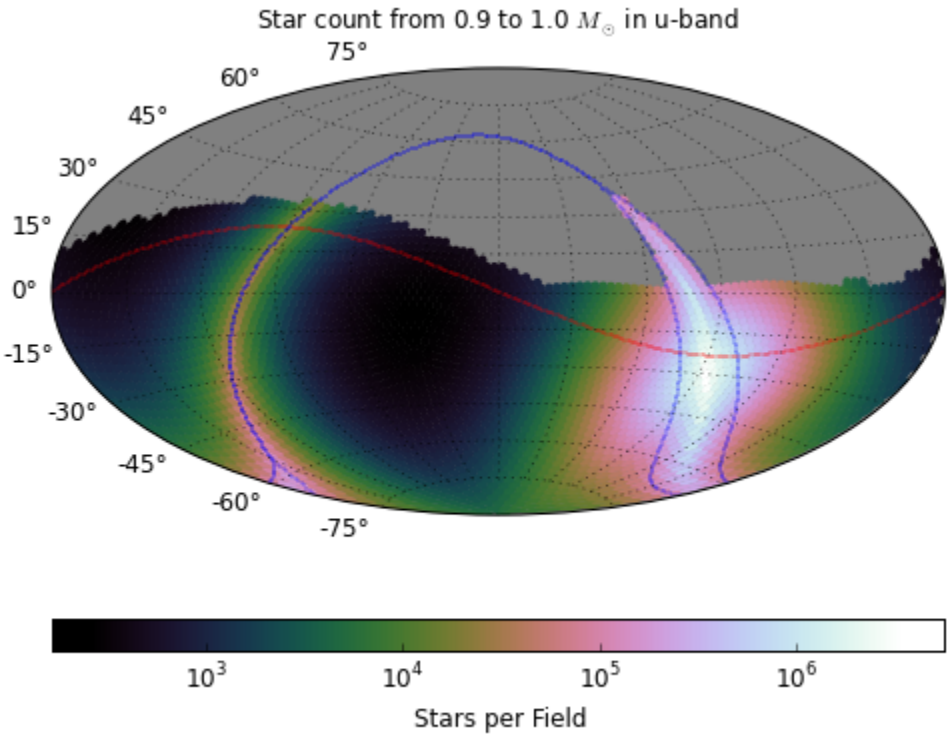

Figure 3.11: Star count per field for the 0.9-1.0 solar mass range and noise less than 0.03 mag. Observations in the *u*-band.

are most useful for finding stellar novae.

The Periodogram Purity Function metric is a tool to probe the phase coverage of the LSST fields. The South Celestrial Pole and the galactic plane subsurveys both have very poor phase coverage and would have very limited usage for studying periodic variables. The Periodogram Purity Function metric currently only uses a single value from the Periodogram Purity Function, and there may be some benefit in an effort to try to better characterize the periodogram purity function by using different criteria to pick this single value. Additionally, while a larger periodogram purity function value should result in the period being easier to detect and characterize correctly, it is not readily apparent to what extent an increase in this value will result in an increased likelihood of detection.

In the Field Star Count metric, we provide a tool to quantify the number of stellar sources that will be present in LSST fields. We look at a few limited mass ranges in particular bands, but it provides indications that we can reasonably expect the number of stars around which planets could be detected (not saturated but sufficiently low noise) is a number in the tens of thousands per field for red dwarfs, and for solar-mass stars exceeds 100,000 stars per field in the direction of the galactic bulge. To use the example of transiting planets where hot Jupiters occur around 1% of stars, and around 10% of hot Jupiters transit, this represents around 10 transiting hot Jupiters around red dwarfs per field (30,000 total), and exceeding 100 transiting hot Jupiters on 0.9-1.0 $M_{\odot}$ stars per field in the direction of the galactic bulge. This doesn't include detection efficiencies or other regimes of planets, but provides a good example of the information that can be gained from considering the number of stellar sources in LSST.

More broadly, the Metric Analysis Framework that was created for LSST can be used for other surveys with a suitable OpSim file. The Observational Triplets and Periodogram Purity Function metrics are writen with only a consideration for the time of observation and the band observed, neither of which are parameters unique to LSST and so these metrics can be applied directly to other surveys using MAF. The Field Star Count metric can, in principle, also be applied to any other survey that MAF can be used with. The current writing of this metric, however, does use an LSST specific field size and noise model, so these would need to be adapted to use the Field Star Count metric with a different survey. As the metric code is all publicly accessable, this would be a relatively simple adjustment to be compatable with a new survey. This adaptability for MAF allows for these metrics to be used in cadence analysis for other surveys in development beyond LSST.

Chapter 4

The Stability of F-star Brightness on Century Timescales

4.1 Introduction

KIC 8462852 (TYC 3162-665-1, and eponymously nicknamed "Boyajian's Star") is an F3 V/IV star in the *Kepler* field. It was discovered to undergo aperiodic, irregular dimming events, with reductions in flux as large as 20% [Boyajian et al., 2016]. The star does not have any apparent infrared excess that would indicate the presence of a large amount of circumstellar dust [Marengo et al., 2015, Thompson et al., 2016]. Thus, one explanation that has been put forward is a family of comets orbiting the star [Boyajian et al., 2016, Bodman and Quillen, 2016]. More exotic solutions have also been proposed, such as that KIC 8462852 is host to a Dyson swarm [Wright et al., 2015], and SETI observations have produced null results in searches for signals in optical and radio wavelengths [Schuetz et al., 2016, Harp et al., 2016].

The time baseline provided by *Kepler* was expanded greatly by the incorporation of DASCH (Digital Access to a Sky Century@Harvard) archival data over the past century in an initial analysis carried out by Schaefer [2016]. That analysis provided comparisons between the light curve of KIC 8462852 and those of two similar stars in close proximity to KIC 8462852 in the photographic plates. Schaefer [2016] finds that KIC 8462852 has steadily dimmed at a rate of 0.165 mag/century over the entire DASCH light curve, which could favor alternatives to the comet hypothesis. In contrast, a follow-up analysis by Hippke et al. [2016], which used a larger sample of ~50 DASCH comparison stars, finds that KIC 8462852 is not unique and that many DASCH stars show apparent long-term trends in their light curves. Those authors suggest that many of the stars experience structural breaks in their light curves, likely due to systematics, and are better modeled using two linear trends with a discontinuity at some year. In particular, the so-called "Menzel

gap" in the 1950s and 1960s during which few observations were obtained, may coincide with a systematic offset in the full set of DASCH light curves that is responsible for the apparent long-term trends in the light curves of KIC 8462852 and many other stars.

In this paper[1], we aim to more comprehensively assess the stability of brightness on century timescales for a large sample of F stars in order to understand the phenomenon of long-timescale variations generally as well as to specifically fully assess the DASCH photometry for KIC 8462852 to determine the significance of any long-term trends. In Section 2, we present the DASCH photometry and the large set of comparison stars we have identified. In Section 3, we examine the long-term trend of KIC 8462852 as well as several hundred other comparable stars to establish the degree to which KIC 8462852 is typical of DASCH stars in terms of the apparent long-timescale trend in its light curve. We also examine if a systematic break at a single year can reduce the number of stars whose light curves appear to have significant long-term trends. Finally, in Section 4, we highlight a sample of stars that may have real long-term trends in their light curves after the likelihood of such trends have been ruled out for KIC 8462852 and for most other comparable stars. We conclude in Section 5 with a summary of our conclusions.

4.2 Data and Methods

4.2.1 Digital Access to a Sky Century@Harvard (DASCH) project Light Curves

The Digital Access to a Sky Century@Harvard (DASCH) project is on ongoing project to digitize ~500,000 glass photographic plates at Harvard College Observatory from observations with several telescopes from 1880 to 1985. For details of the image digitation and photometric extraction process, including parameters for measuring the stellar images, telescope image scale, etc., we refer the reader to prior DASCH publications [Simcoe et al., 2006, Grindlay et al., 2009, Laycock et al., 2010, Grindlay et al., 2012, Tang et al., 2013a].

[1]This chapter was originally published as [Lund et al., 2016a]

Briefly, these plates are scanned, and a data reduction pipeline is used to identify sources, calibrate each plate, generate light curves, and flag any potentially bad data points [Laycock et al., 2010]. To calculate magnitudes, source stars are compared to catalog stars, and DASCH has used three different stellar catalogs for this purpose. For a thorough overview of the process that DASCH implemented for photometric calibration, including performance relative to standard stars, see Laycock et al. [2010]. The initial photometry was done by using the Hubble Guide Star catalogue (GSC2.3) B-band magnitudes to calibrate the photometry, with precision of ±0.2 mag over a century [Grindlay et al., 2012]. This is being improved by also using the all-sky APASS CCD survey, which allows for improved photometry and color-corrections, improving the precision to ± 0.1 mag over a century [Grindlay et al., 2012]. Currently these light curves are available for all sources with galactic latitudes greater than 30. Additionally, the DASCH team has released light curves for the Kepler field, calibrating this data using the *Kepler* Input Catalog (KIC) *g*-magnitudes. The light curves produced for the Kepler field have a precision of ∼0.1 mag [Tang et al., 2013a].

We have collected light curves from the DASCH website (http://dasch.rc.fas.harvard.edu/)[2] . To maintain high data quality, from every light curve we remove all points that a non-zero value for AFLAGS (defined by the DASCH interface as "'fatal' photometry pipeline errors"), and only analyze light curves that have at least 100 points after this cleaning of the data. This criterion for removing poor quality data is the same that was used by the DASCH team in Tang et al. [2013b]. We use light curves from both the APASS B-band and KIC *g*-band calibrations.

[2]Different formats for obtaining light curves from the DASCH web site can result in variations in the light curve that is returned, and so we outline the procedure for obtaining the light curves here. The DASCH light curves are accessible through http://dasch.rc.fas.harvard.edu/lightcurve.php. Stars were queried in batches of 10, using the default settings for minimum number of measurements (1), distance in arcseconds (5), and using a frame format. From the upper-left frame of the next window, we then selected "Download all points in table form." This provides a list of the results for all sources in the query. The downloaded files are the short form Starbase (tab-delimited ASCII) tables, also listed as option 'A' for each source.

4.2.2 Selection of F-Star Study Sample

We selected two samples of F-type stars based on the properties of KIC 8462842. From the Tycho-2 Spectral Type Catalog [Wright et al., 2003], we selected all F3 IV and V stars that are also included in the DASCH main data releases. The resulting list has 967 stars for which we gathered light curves from the DASCH database. These stars are predominantly between 9th and 11th magnitude and calibrated to the APASS B-band. KIC 8462852, for instance, is $B = 12.8$.

The second sample of stars was chosen from the Kepler field based on the *Kepler* Input Catalog [Brown et al., 2011]. We selected stars that were similar to KIC 8462852, and so limited the sample to KIC stars with effective temperatures within 100K of KIC 8462852, as well as within 5% in stellar radius and 10% in log(g). The KIC lists KIC 8462852 as having a temperature of 6584K, a log(g) of 4.124, and a stellar radius of 1.699 $R_{\odot}$. The resulting list has 559 stars, and we gathered light curves for them the DASCH database, calibrated to the KIC g-band.

For all stars from both catalogs we apply the cuts to the light curves discussed in §4.2.1. We then apply a final cut to our lists of stars, requiring all stars to have at least 100 remaining data points, and are left with 337 stars from the APASS calibration, and 307 stars from the KIC calibration, for a total of 644 light curves.

4.3 Characterizing Long-Term Trends in the Light Curves

In this section, we describe the statistical tests that we employ in our analysis of long-term trends in the light curves, with particular attention to assessing potential abrupt systematics ("structural breaks") in the light curves. It has been proposed [Hippke et al., 2016] that many DASCH light curves exhibit structural breaks at a given point in time due to a gap in observations. This occurs during the so-called Menzel gap in the 1950s and 1960s, during which few observations were obtained, and appears in all light curves. We examine

that claim in more detail in the section, and apply several statistical tests to further evaluate that claim by examining the light curves we have from both the APASS calibrated and KIC calibrated stars.

If we presume that the light curves of stars in the sample are on average flat over a century of observations, then we should expect that the t-statistics for the slopes follow a normal distribution. In other words, we expect that just by the nature of noise in the data, we should find ~ 200 stars (of over 600 stars examined) with slopes greater than 1 σ from 0, ~ 30 stars with slopes greater than 2 σ from 0, and only 1 or 2 stars with slopes greater than 3 σ from 0. However, a systematic structural break in the light curves will cause many more apparent outliers than these expectation values. In contrast to those expecations, in our data set there are 189 stars with slopes greater than 3 σ from 0, 50 of which are greater than 5 σ from 0. In this section, we try three different statistical approaches to test for a systematic break in the data.

4.3.1 Statistical Search for Systematic Break

We begin our hypothesis with the possibility of a systematic structural break in the data, such that at some year the best-fit lines for all the stars is discontinuous. To determine the best year to place this break, we iterate over 1902 thrugh 1987; for each year, we split every light curve at that year, fit each of the segments to a line, and calculate the z-score for the significance of the slopes. (This approach was taken to explicitly test for any possible common points for the break, rather than automatically assuming the break happens at the Menzel gap.) If the true slope for every line is near zero, then the z-scores would follow a normal distribution. We then use a Pearson's χ^2 test [Pearson, 1900] to compare the distribution of the z-score of the slopes to a normal distribution with the intention of finding the year where the z-scores are closest to a normal distribution. This was done by calculating the reduced χ^2 for this distribution, shown in Fig 4.1 and treating the Tycho and KIC stellar samples independently. The binned data for the full light curves and for the

pre-1970 light curves are both shown in Fig 4.2.

We find that the optimal break using the light curves based on the APASS calibration is 1954, and the optimal break using the light curves based on the KIC calibration is 1971. More significantly, however, we see that in both figures the region where the reduced χ^2 is near its minimum is approximately the same time frame, and this roughly corresponds to the Menzel gap, which spans 1953 to 1969 [Schaefer, 2016]. One prefers the early edge of the Menzel gap while the other prefers the late edge, but because it is a gap in the data these are essentially equivalent choices. Therefore, we use 1970 as a break year going forward for both sets of stars.

Fig 4.2 makes it readily apparent that by splitting the light curves at 1970 and fitting the segments before and after this year independently, we have a distribution that is much more consistent with most stars having slopes of 0 and that many of the large slopes also have large uncertainties on the slopes and do not represent significant detections of non-zero slopes. Note that while the slopes derived from the post-1970 data have larger uncertainties in units of magnitude relative to the pre-1970 data because of the smaller time baseline (~18-y versus ~70-y baseline), the z-score histograms account for this by presenting the comparison in sigma units.

4.3.2 Magnitude Changes Before and After Menzel Gap

We can also take a complementary approach by examining each light curve at three time intervals: 1920-1925, 1945-1950, and 1970-1975. We use a Two Sample Weighted T-Test to compare the two 5-year intervals; the Two Sample Weighted T-Test is used to test the null hypothesis that two samples are drawn from the same underlying distribution [Welch, 1947]. If a star has the same underlying brightness in each 5-year interval that we examined, then our expectation is that the values for the Two Sample Weighted T-Test should not be significantly different for the two intervals.

We find that the distribution of p-values for the second interval (comparing 1945-1950

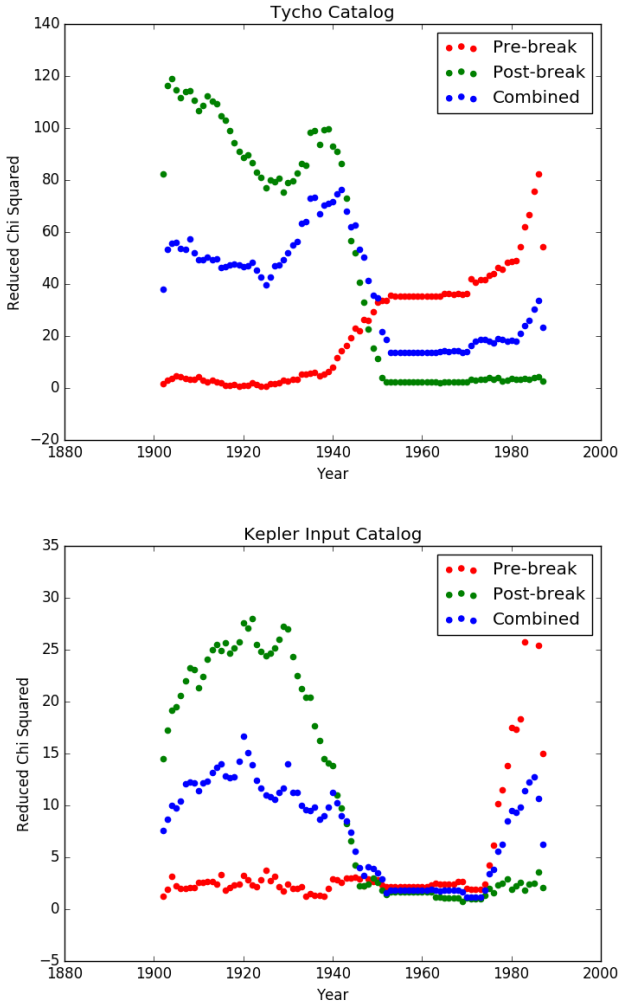

Figure 4.1: Reduced χ^2 values of the t-statistic of the slope as compared to a normal distribution and as a function of the year at which the light curves are split. The red points are the comparison to a normal distribution only for slopes prior to the break, the green points are only slopes after the break, and the blue points are the reduced χ^2 for the slopes before and after the break jointly. The Tycho catalog (top) reaches a minimum in 1954, and the KIC (bottom) reaches a minimum in 1971. The Menzel gap is generally defined as being from 1953 to 1969.

and 1970-1975) is overall higher than the distribution for the first interval (comparing 1920-1925 and 1945-1950), as shown in Fig 4.3, and according to a k-sample Anderson-Darling test we can reject the hypothesis that the distribution of p-values are the same with 92% significance [Anderson and Darling, 1954]. This indicates that data from immediately

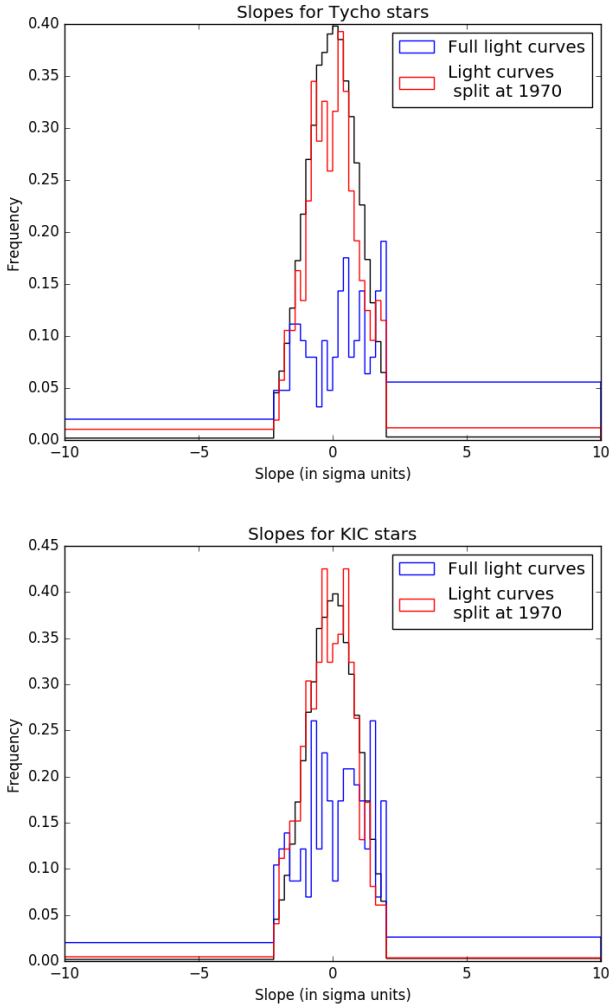

Figure 4.2: Histograms of the z-scores for the slopes of all stars in the Tycho catalog (top) and KIC (bottom). The black line marks the histogram of a normal distribution. The blue line is the distribution of z-scores of the slopes when we fit the entire light curve at once. The red line is the distribution of the z-scores of the slopes when we split each light curve in two pieces at 1970 before fitting lines to the light curves.

before and after the Menzel gap are less similar than data separated by a similar interval but part of the continuous pre-Menzel Gap data.

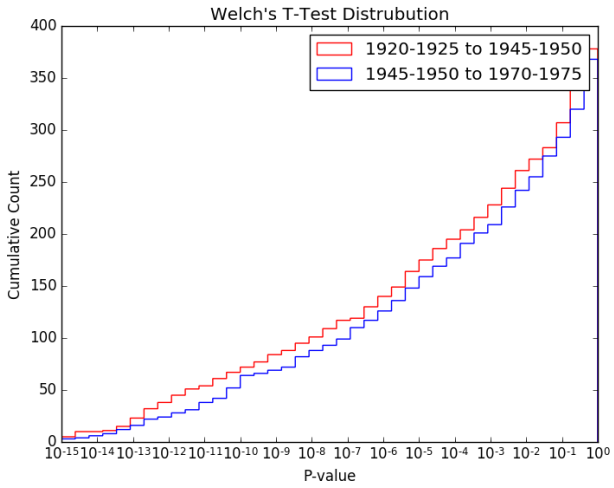

Figure 4.3: Welch's t-test distributions for light curves. The red line represents the first interval (comparing 1920-1925 and 1945-1950), and the blue line represents the second interval (comparing 1945-1950 and 1970-1975).

4.3.3 Slope changes Before and After Menzel Gap

Finally, we used a Kendall tau test [Kendall, 1938] to assess the likelihood of slope changes in the data before and after the Menzel gap. The Kendall tau test is a non parametric test that compares the ordinal association in two quantities. For this case, we compared the slopes of the light curves before and after 1970. If single lines are suitable fits to the data across the full time span, there should be a high correlation in the slopes of the light curves. In other words, the stars with the highest slopes before 1970 should also have the highest slopes after 1970, and vice versa. In the Kendall tau test, a tau statistic value of 1 represents perfect correlation, 0 represents no correlation, and -1 represents perfect anti-correlation. For the APASS calibrated light curves we find a tau statistic of -0.058, and for the KIC calibrated light curves, we find a tau statistic of -0.060. This again suggests that the data before and after the Menzel gap cannot be combined by fitting a single line, since these results are close to 0 rather than 1.

4.3.4 Assessing Structural Breaks in the Light Curves: The Menzel Gap

In this section, we have approached the existence of a structural break through multiple methods. We have found that the year range that is most indicative of a systematic break is coincident with the Menzel Gap. As an ensemble, the light curves also undergo more significant changes before and after the Menzel Gap than they did over a similar time span earlier in the light curves. Finally, there's no correlation between the slopes before 1970 and the slopes after 1970, as would be expected if the data could best be fit by a single line. These considerations all suggest that the data possess a systematic break during the Menzel Gap, which includes the year 1970. We thus assess all of the light curves in our following analysis with this consideration in mind, and by using only pre-1970 data.

4.4 Results: Analysis of Long-term Photometric Trends

4.4.1 F Stars

We start with the two lists of stars we discussed in §4.2.2. We calculate the slope of a line using the method of weighted least squares for the light curves of all 644 stars, as shown in Fig 4.4. Since there is no astrophysical reason to expect that long-term brightness changes would be present as a function of apparent magnitude for a set of dwarf stars, the trends towards brightening that we see for both bright (≤ 9) stars and faint (≥ 13.5) stars are likely indicative of systematics at the bright and faint limits of the calibration, respectively. Within these magnitude limits, KIC 8462852 is not unique in having a significant slope, and a large number of stars show slopes that deviate from a flat slope by at least 1 σ. Even at approximately the same magnitude, there are several stars with more significant slopes, either dimming or brightening.

We reassess the distribution of slopes as a function of magnitude in Fig 4.5, where we have now only used the pre-1970 data in the weighted least-squares linear fit. The typical errors of the slope values are slightly larger, but the brightening trends seen for bright and

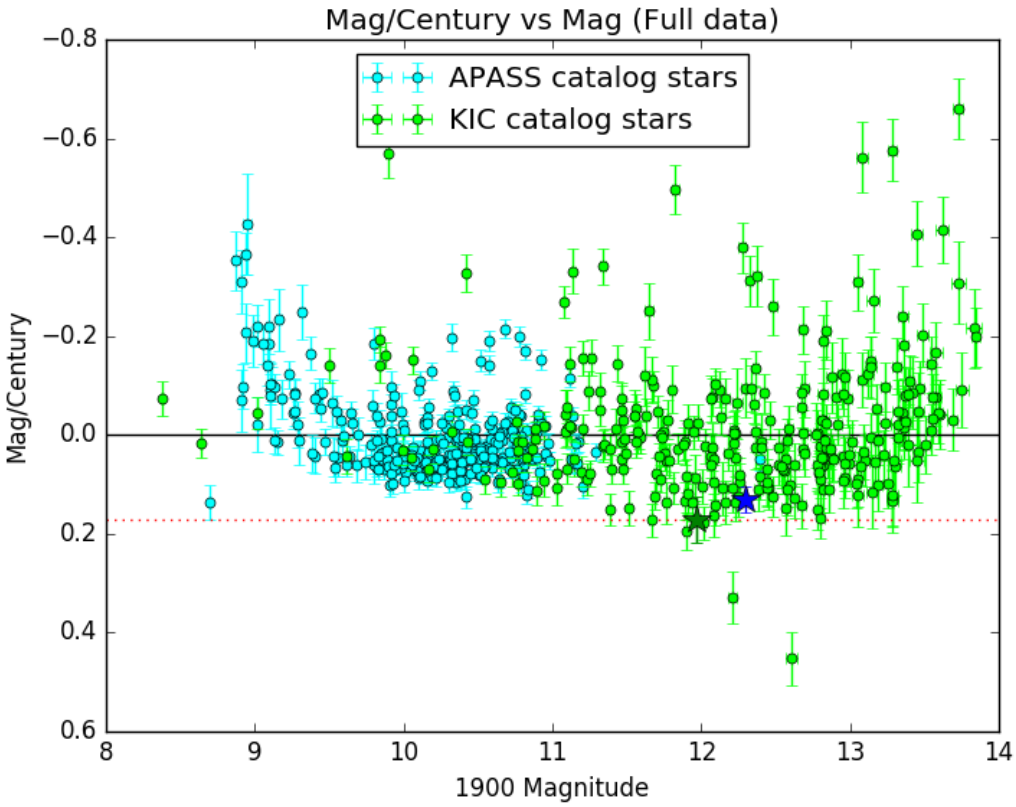

Figure 4.4: Century-long slopes as a function of magnitude for all stellar twins of KIC 8462852. Cyan circles represent light curves using the Tycho catalog calibrated to the APASS B-band. Green circles represent light curves using the KIC calibrated to the g-band. The dark blue and dark green star symbols are KIC 8462852 calibrated to the APASS B-band and KIC g-band, respectively. The red dotted line highlights the slope value of the light curve of KIC 8462852 from the KIC calibration.

faint stars are less noticeable, since the median of the ensemble of slopes is closer to 0, and the scatter at the bright and faint ends is no longer as asymmetric. Additionally, the values for KIC 8462852 are now consistent with a slope of 0 (the slopes for the post-1970 data in KIC 8462852 are also displayed in Fig 4.6).

4.4.2 KIC 8462852

We first examine two light curves for KIC 8462852, one for each of the calibration catalogs that we use. The APASS B-band calibrated light curve consists of 1887 points,

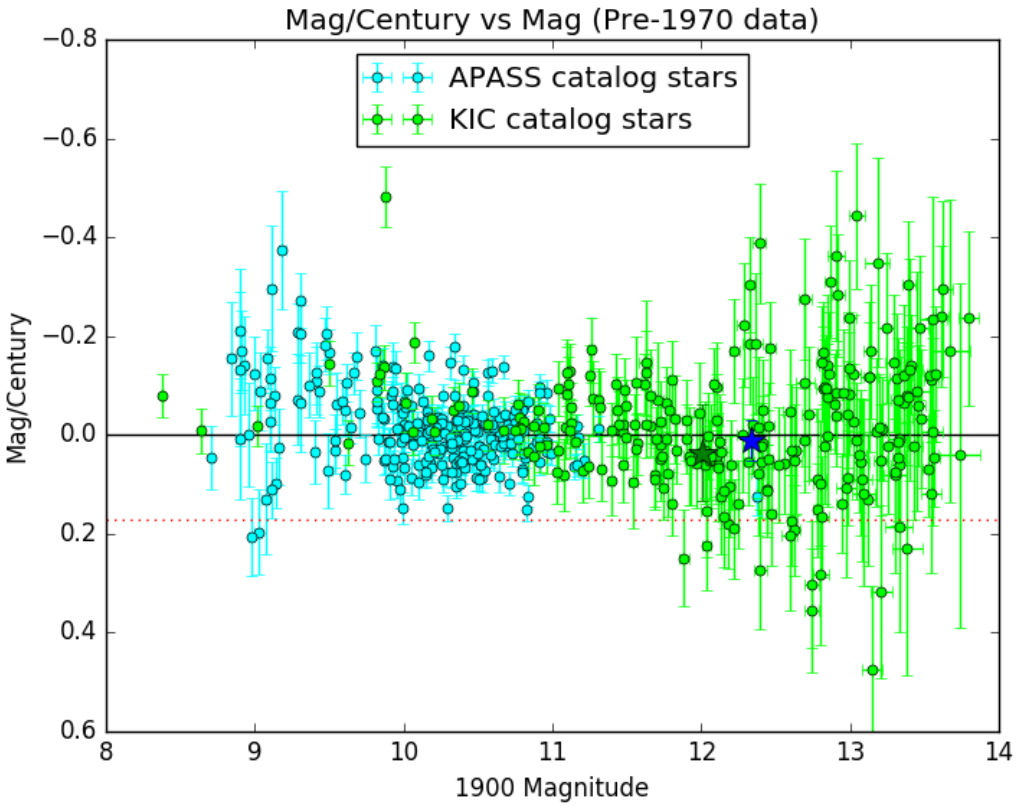

Figure 4.5: Slopes as a function of magnitude for all stellar twins of KIC 8462852, using only the pre-1970 light curves. Cyan circles represent light curves using the Tycho catalog calibrated to the APASS B-band. Green circles represent light curves using the KIC calibrated to the g-band. The dark blue and dark green star symbols are KIC 8462852 calibrated to the APASS B-band and KIC g-band, respectively. The red dotted line highlights the slope value of the light curve of KIC 8462852 from the KIC *g*-band for the full data set.

which is reduced to 781 points by our quality controls discussed in §4.2.1. The KIC g-band calibrated light curve consists of 1252 points, which is reduced to 575 points by our quality controls. In both cases, we use the method of weighted least squares to fit a single line to the full light curves, shown in Fig 4.6. In both cases, we find that fitting a single line to the entire data results in a non-zero slope with at least 3 σ significance, with the slope of the APASS calibrated light curve equal to 0.129 ± 0.026 mag/century, and the slope of the KIC calibrated light curve equal to 0.171 ± 0.047 mag/century.

We also break the light curves into two segments that we fit independently, before and

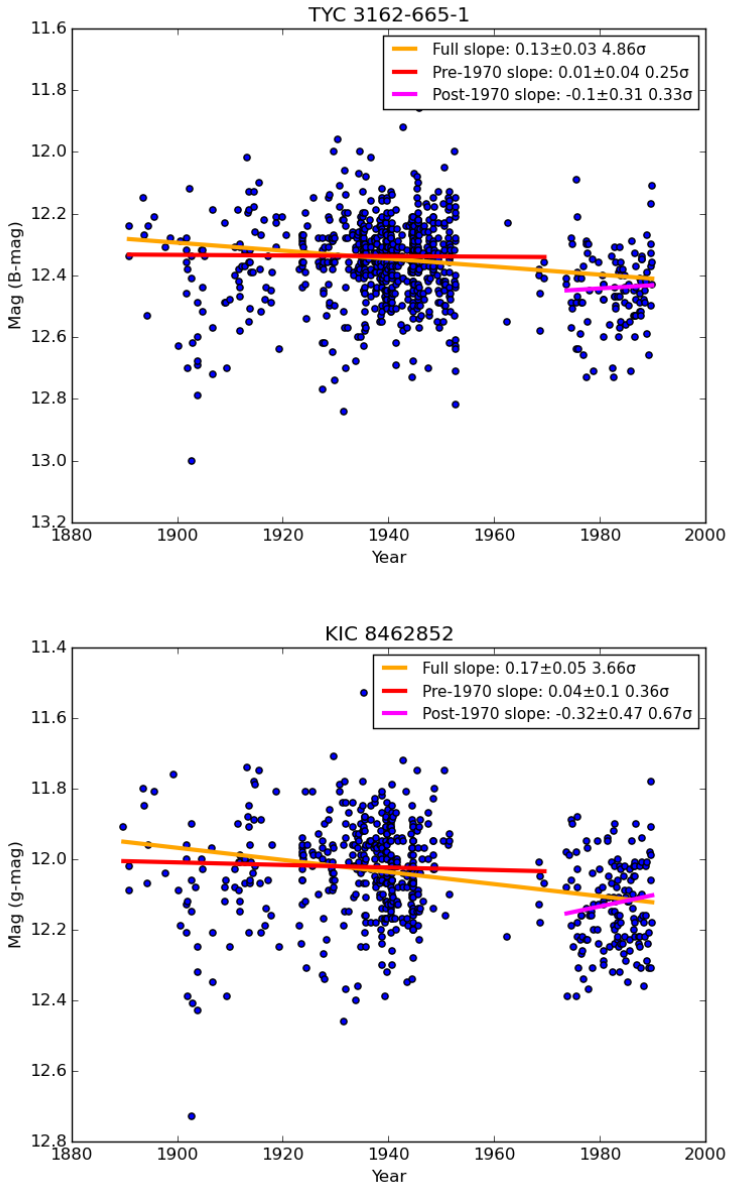

Figure 4.6: KIC 8462852 light curves in the APASS B-band calibration (top) and KIC g-band calibration (bottom), with data cuts as described in §4.2.1. A linear fit to the entire light curve results in a dimming trend significant to at least 3 sigma. When the light curves are divided into two segments, before and after 1970, the linear fits for both segments are consistent with no slope.

after 1970; this choice of year explained in §4.3. When the light curves are broken at this year and using a critical p-value of 0.05, we cannot reject the hypothesis that the data is a flat line for any of the line segments. For both of the light curves we have also used a Two

Sample Weighted T-Test to test if the magnitudes before and after 1970 are drawn from the same population. We are able to reject the hypothesis that pre-1970 and post-1970 data are drawn from the same distribution with a critical p-value of 0.05. These results suggest that the century-long slope we observe is a result of an abrupt change in the magnitude around 1970 rather than a gradual and consistent dimming over the course of a century.

4.5 Discussion

The DASCH light curves for KIC 8462852, when examined on their own and with a single linear fit, suggest that the star has undergone a century-long dimming event that would be particularly noteworthy given the events in the *Kepler* light curve that were announced by Boyajian et al. [2016]. Trends in the DASCH light curve of KIC 8462852, however, can also be explained by a structural break in an otherwise flat light curve. When light curves of a set of stars with physical properties similar to KIC 8462852 are split at 1970, we find that the t-statistics of the slopes are much closer to a normal distribution, as would be expected if all light curves were truly flat.

We do still find an excess of stars with significant slopes, with 9 stars from the APASS B-band calibration and 2 stars from the KIC g-band calibration having slopes that are non-zero to greater than 5 σ. All of these stars are listed in Table 4.1, and the light curves are included in Appendix A. These stars can be accounted for with three possible explanations: (1) there are additional structural breaks beyond the Menzel Gap that have not been accounted for; (2) systematics such as slightly incorrect error estimations result in underestimating the error in the slope; and (3) some of these stars may actually have changed brightness over the last century.

If any of these trends are real astrophysical signals, it would indicate that a small fraction of all F stars can vary on century-long timescales. However it would need to be determined that the brightening or dimming is coming from the observed stars, and not from other nearby stars.

Table 4.1: Stars with Significant (5σ) Long-term Photometric Trends

Star	Full Slope	Pre-1970 Slope	Post-1970 Slope
TYC-6727-00524-1	0.121±0.017	0.150±0.025	0.076±0.258
TYC-6174-00949-1	-0.215±0.020	-0.138±0.024	-0.264±0.340
TYC-5531-01038-1	-0.169±0.016	-0.144±0.021	-0.321±0.270
TYC-5554-01593-1	-0.023±0.019	-0.161±0.031	0.120±0.263
TYC-6178-00821-1	-0.088±0.020	-0.178±0.028	-0.213±0.277
TYC-6749-00508-1	0.126±0.021	0.141±0.025	0.108±0.411
TYC-6160-00274-1	0.111±0.017	0.124±0.022	0.158±0.325
TYC-5554-01017-1	-0.022±0.017	-0.130±0.026	0.603±0.275
TYC-6165-01434-1	0.096±0.021	0.149±0.027	-0.198±0.409
K3868420	-0.571±0.051	-0.481±0.061	1.101±1.422
K11802860	0.452±0.055	1.090±0.180	-1.282±0.487

4.6 Conclusion

The DASCH survey is extremely useful for looking for large changes in flux that have taken place with a century-baseline for a range of sources, including long-period eclipsing binaries [Rodriguez et al., 2016], large-amplitude variation in K giants [Tang et al., 2010], and quasars/blazars [Grindlay et al., 2012]. There are, however, limitations resulting from the lowest photometric errors still being on the order of 0.1 mag, and this results in circumstances that allow for systematics to be mistaken for astrophysical signals. The analysis carried out here indicates that KIC 8462852 is such a case, and that KIC 8462852 itself has not undergone any long-term trends in brightness.

We do find that there are more stars with significant slopes than would be expected if F stars maintained constant luminosity. We include this list of stars to allow for further investigation to determine if systematics may explain these slopes, or if these stars are undergoing significant and unexpected changes in magnitude on long time scales. Examining the high-precision light curves of these stars that may become available with TESS and possibly other surveys, may prove particularly valuable.

Chapter 5

Future Work

5.1 Work in Progress

In this section I briefly outline the current work that is being undertaken as part of improving the characterization of LSST, particularly with a focus on exoplanets. The initial work discussed in Chapter 2 has been expanded to focus on the Large Magellanic Cloud (LMC) specifically, as discussed in 5.1.1. This work will be particularly relevant as the observing cadence for the LMC has yet to be decided, and there is currently a push to have the Magellanic Clouds observed at the high cadence of a deep-drilling field [Szkody et al., 2011]. The second work under development is a compilation of additional metrics that have been developed since the publication of the first set of LSST metrics (Chapter 3). One of these metrics, a metric applied for microlensing with the particular interest of examining LSST's ability to be used for detection of microlensing events, is discussed in 5.1.2.

5.1.1 Transiting Planets in the Large Magellanic Cloud

One of the particularly unique stellar populations that LSST will be able to observe is approximately stellar-mass stars in the LMC, and discoveries of transiting planets here allow not just for the discovery of potentially the first extragalactic planets, but also provide some insight into how the exoplanet occurrence rate varies between stellar populations, particularly as a function of metalicity. We have previously demonstrated the ability of BLS to correctly recover the period of transiting exoplanets at the LMC, however only in the context of deep-drilling fields, and so we only work with the deep-drilling cadence here [Lund et al., 2015].

We begin with a stellar population of 500,000 stars that are of approximately solar-mass (0.9 to 1.1 $M_{\odot}$) with a distribution consistent with a standard Salpeter IMF. We then popu-

late each of these planets using the frequency of planets as a function of radius and period that has been calculated using the *Kepler* sample [Youdin, 2011], and then calculating the geometric probability that each planet transits its host star. From our initial population of stars, we are left with approximately 20,000 transiting planets. We then simulate the six-band light curves for all host stars with transiting planets, and subsequently search for periods using the BLS algorithm. We consider any period within 0.1% of the input period of a transiting planet to be consistent with recovering the transiting planet signal. We provide one such example of a Hot Jupiter at approximately 20 $R_\oplus$ that we successfully recover in Figure 5.1. In this figure, we see that the transiting event is visible by eye when we plot

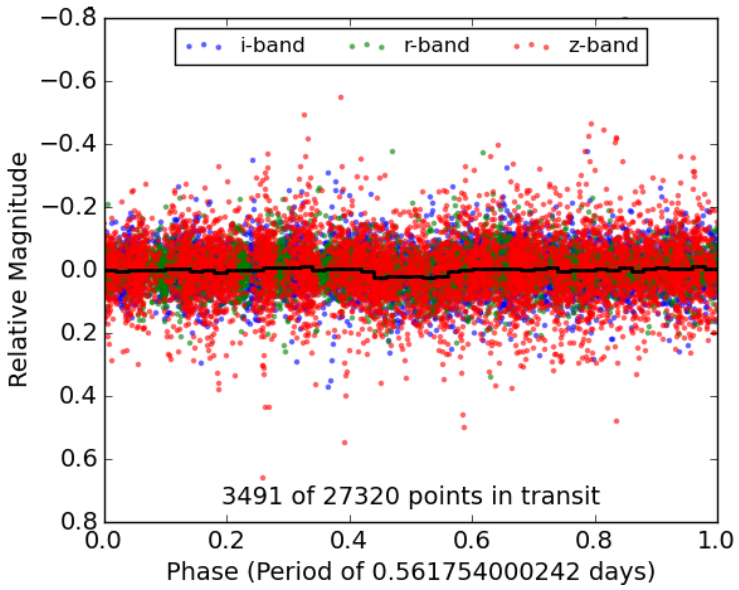

Figure 5.1: A simulated light curve of a transiting Hot Jupiter around a solar-like star at the distance of the Large Magellanic Cloud.

the phased light curve for the star and bin this data. If we look at the entire population of transiting planets, we find that there are a substantial number of planets whose periods are successfully recovered, as shown by the red points in Figure 5.2. These recoveries occur in two regimes; the first regime is that we see a large portion of Hot Jupiters (planets with radii greater than 10 $R_\oplus$ and periods under 10 days), and represent what we believe

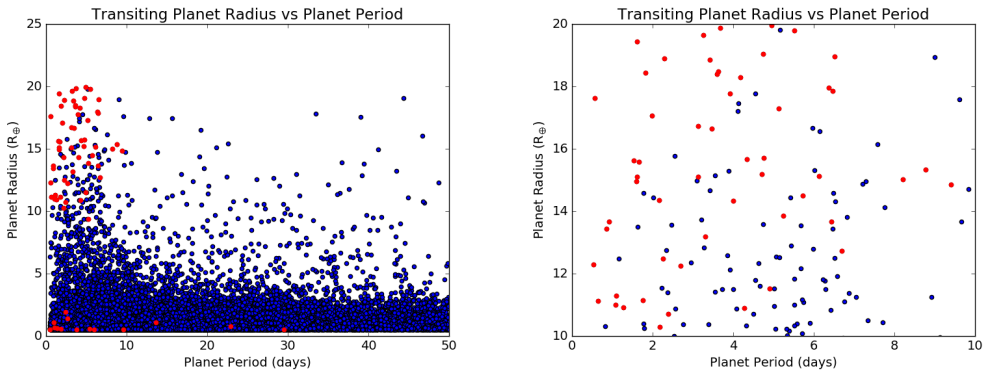

Figure 5.2: All transiting planets in the LMC simulations are shown in blue with those planets whose periods are accurately recovered using BLS shown in red (left). The figure at right zooms in specifically on the Hot Jupiter regieme, which is where non-spurious recoveries are concentrated.

to be robust period recoveries. The second area where we see planets that are recovered are of exoplanets that are of approximately earth radius, and it is important to note that the smallest planets are also expected to be the most common, and so the exoplanet period recoveries at small periods should be interpreted as spurious recoveries where the top BLS period happened to be consistent with the input period of a planet, rather than actual detections of exoplanet signals. It should also be noted that while there is insight provided by determining which exoplanets could be properly recovered, the true goal will be to calculate the number of exoplanets that we can detect (i.e. what number of exoplanets could be found without a priori knowledge of their input periods).

This work is still undergoing a more rigorous analysis before it will be complete, however we can provide an optimistic rough estimate based off the initial simulations. If we can effectively separate out all Hot Jupiters that we are correctly able to recover the periods of from any statistical false positives that may occur, then based off of our period recoverability work and a planet occurrence rate similar to the *Kepler* field, then we would expect to find 1 transiting planet per 10,000 solar-type stars in the LMC. This may be interpreted as an overly optimistic characterization, however it also provides an initial hypothesis. Testing

this hypothesis wil allow us to see how planet formation rates actually differ in the LMC, provided that LSST does choose to observe the LMC at the deep-drilling cadence.

5.1.2 Microlensing Metric for LSST

LSST's application toward microlensing events can be, broadly, categorized into two different approaches. The first approach is to attempt to use LSST to provide a solution for a microlensing event as an independent set of observations, and this will be challenging given the sparse nature of the data. The second approach, however, is to use LSST alerts as triggers for microlensing follow-up observations by dedicated microlensing telescopes, in much the same way that microlensing searches such as OGLE and MOA already work. in this latter case, the interest is simply to have a microlensing event be observed early enough by LSST that most of the event can be thoroughly characterized by follow-up observations. To examine how well LSST will be able to provide these alerts, we introduce a simple metric towards this goal.

To define this metric, we set up a simplified set of criteria for detecting a microlensing event. We require that we have three LSST observations (presently, with no band restrictions) that all occur within some maximum window of time, and where subsequent events are separated by some minimum window of time. The precise values of these intervals will depend on the characteristic time scale of the event we are interested in, and so for now we examine the signals that we would expect for a lone planetary microlensing event as well as for a standard stellar microlensing event. The metric itself displays what fraction of time a microlensing event would then trigger additional followup under this criteria. For a planetary signal, we require that the observations have a minimum spacing of 1 hour and all occur over a 2-day window, and the results are shown in Figure 5.3. We see in these results that for planetary microlensing, the cadence is very poorly suited to this, with the exception of the deep-drilling fields at high cadence. The time scales of stellar microlensing events is longer, and so here we require observations that are separated by at least 2 days, but all

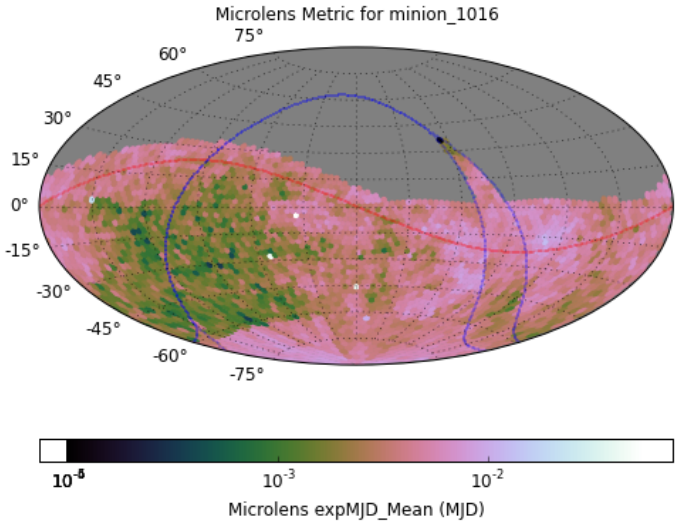

Figure 5.3: Microlensing metric results for a short timescale event consistent with a planetary microlensing signal.

take place in a 7-day window, with results in Figure 5.4. In general, these events are more

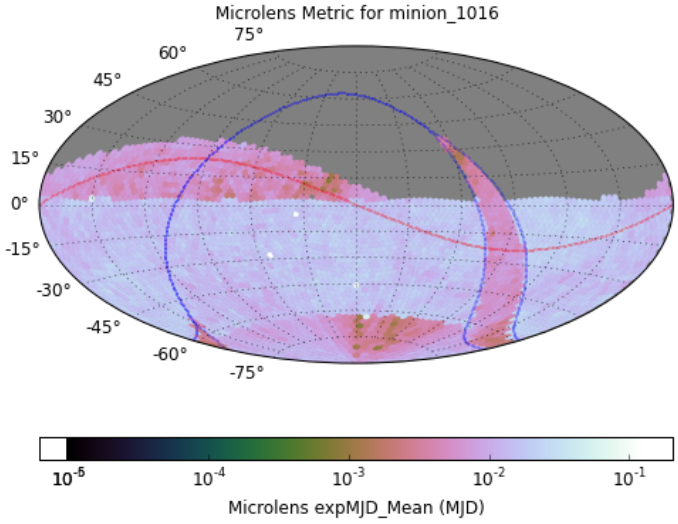

Figure 5.4: Microlensing metric results for a medium timescale event consistent with a stellar microlensing signal.

likely to fit our criteria to trigger an alert. More interestingly, however, is the initial observation that the LSST cadence is actually more favorable to detecting stellar microlensing events outside the plane than in it. This may be rather concerning, as the plane and Glactic Bulge are precisely where these events are most likely to occur, and suggests that it will be worth considering changes to the cadence to improve performance in the Galactic Plane. An important caveat in interpretting these metric results is that while this metric is returning the fraction of time where a microlensing event could occur and be detected, a more thorough assessment of LSST's microlensing capability should combine this metric with the frequency of a microlensing event occurring as a function of area of the sky, and that parameter of microlensing frequency is not in any currently existing metric.

5.2 Conclusion

To understand the scientific contributions that can be made by any survey, there is a need to also understand the limitations and characteristics of that survey. Much of the focus in this work has been specifically working with LSST, and even more specifically, the contributions that LSST can make to exoplanet science by discovering exoplanets around host stars in various stellar populations. The simple capacity for LSST to observe transit events from exoplanets is demonstrated in Lund et al. [2015], and this framework was also used by Jacklin et al. [2015]. Currently in development is additional work on exoplanet yields for the Large Magellanic Cloud, which will provide better constraints on the number of extragalactic planets that may be obtained with LSST. Additional future work will be calculating transiting exoplanet yields for additional stellar populations within LSST, and expanding this search to include exoplanets that can be detected using microlensing, or in combination with other surveys. In the process, this work has also involved developing tools to help characterize LSST by providing metrics to the community, particularly for exploring how LSST will perform for variable sources [Lund et al., 2016b].

These questions with LSST, however, speak to a more general need to understand the

limitations of surveys. As future work with LSST will, no doubt, require better understanding of the sparsely-sampled nature of the data, the lessons that are learned to find planets in LSST will also be applicable to other instruments that are providing similarly-sparse light curves, such as Hipparcos or Gaia. The most fundamental questions, such as understanding the limits in interpreting the data, also applies to archival data sets that don't represent ongoing surveys, such as the work in Lund et al. [2016a]. The era of big data provides a rich opportunity for new science to be conducted, but to have confidence in these results it will be important to understand not just the strengths but also the weaknesses and limitations of the myriad of surveys that have been or will be conducted.

BIBLIOGRAPHY

C. W. Allen and A. N. Cox. *Allen's Astrophysical Quantities*. New York: Springer, 4 edition, 2000. ISBN 978-0387987460. doi: 10.1063/1.1325201.

T. W. Anderson and D. A. Darling. A Test of Goodness of Fit. *Journal of the American Statistical Association*, 49(268):765, dec 1954. ISSN 01621459. doi: 10.2307/2281537.

G. Á. Bakos, J. D. Hartman, G. Torres, G. Kovács, R. W. Noyes, D. W. Latham, D. D. Sasselov, and B. Béky. Planets from the HATNet project. 11(August):01002, 2011. ISSN 2100014X. doi: 10.1051/epjconf/20101101002.

Gibor Basri, Lucianne M. Walkowicz, Natalie Batalha, Ronald L. Gilliland, Jon Jenkins, William J. Borucki, David Koch, Doug Caldwell, Andrea K. Dupree, David W. Latham, Geoffrey W. Marcy, Soeren Meibom, and Tim Brown. Photometric Variability in Kepler Target Stars. Ii. an Overview of Amplitude, Periodicity, and Rotation in First Quarter Data. *The Astronomical Journal*, 141(1):20, jan 2011. ISSN 0004-6256. doi: 10.1088/0004-6256/141/1/20.

Natalie M. Batalha, Jason F. Rowe, Stephen T. Bryson, Thomas Barclay, Christopher J. Burke, Douglas a. Caldwell, Jessie L. Christiansen, Fergal Mullally, Susan E. Thompson, Timothy M. Brown, Andrea K. Dupree, Daniel C. Fabrycky, Eric B. Ford, Jonathan J. Fortney, Ronald L. Gilliland, Howard Isaacson, David W. Latham, Geoffrey W. Marcy, Samuel N. Quinn, Darin Ragozzine, Avi Shporer, William J. Borucki, David R. Ciardi, Thomas N. Gautier, Michael R. Haas, Jon M. Jenkins, David G. Koch, Jack J. Lissauer, William Rapin, Gibor S. Basri, Alan P. Boss, Lars A. Buchhave, Joshua A. Carter, David Charbonneau, Joergen Christensen-Dalsgaard, Bruce D. Clarke, William D. Cochran, Brice-Olivier Demory, Jean-Michel Desert, Edna Devore, Laurance R. Doyle, Gilbert A. Esquerdo, Mark Everett, Francois Fressin, John C. Geary, Forrest R. Girouard, Alan

Gould, Jennifer R. Hall, Matthew J. Holman, Andrew W. Howard, Steve B. Howell, Khadeejah A. Ibrahim, Karen Kinemuchi, Hans Kjeldsen, Todd C. Klaus, Jie Li, Philip W. Lucas, Søren Meibom, Robert L. Morris, Andrej Prša, Elisa Quintana, Dwight T. Sanderfer, Dimitar Sasselov, Shawn E. Seader, Jeffrey C. Smith, Jason H. Steffen, Martin Still, Martin C. Stumpe, Jill C. Tarter, Peter Tenenbaum, Guillermo Torres, Joseph D. Twicken, Kamal Uddin, Jeffrey Van Cleve, Lucianne Walkowicz, and William F. Welsh. Planetary Candidates Observed By Kepler . III. Analysis of the First 16 Months of Data. *The Astrophysical Journal Supplement Series*, 204(2):24, feb 2013. ISSN 0067-0049. doi: 10.1088/0067-0049/204/2/24.

Thomas G. Beatty and B. Scott Gaudi. Predicting the Yields of Photometric Surveys for Transiting Extrasolar Planets. *The Astrophysical Journal*, 686(2):1302–1330, oct 2008. ISSN 0004-637X. doi: 10.1086/591441.

Eva H. L. Bodman and Alice Quillen. KIC 8462852: Transit of a Large Comet Family. *The Astrophysical Journal*, 819(2):L34, 2016. ISSN 2041-8213. doi: 10.3847/2041-8205/819/2/L34.

T. S. Boyajian, D. M. LaCourse, S. A. Rappaport, D. Fabrycky, D. A. Fischer, D. Gandolfi, G. M. Kennedy, H. Korhonen, M. C. Liu, A. Moor, K. Olah, K. Vida, M. C. Wyatt, W. M. J. Best, J. Brewer, F. Ciesla, B. Csák, H. J. Deeg, T. J. Dupuy, G. Handler, K. Heng, S. B. Howell, S. T. Ishikawa, J. Kovács, T. Kozakis, L. Kriskovics, J. Lehtinen, C. Lintott, S. Lynn, D. Nespral, S. Nikbakhsh, K. Schawinski, J. R. Schmitt, A. M. Smith, Gy. Szabo, R. Szabo, J. Viuho, J. Wang, A. Weiksnar, M. Bosch, J. L. Connors, S. Goodman, G. Green, A. J. Hoekstra, T. Jebson, K. J. Jek, M. R. Omohundro, H. M. Schwengeler, and A. Szewczyk. Planet Hunters IX. KIC 8462852 – where's the flux? *Monthly Notices of the Royal Astronomical Society*, 457(4):3988–4004, apr 2016. ISSN 0035-8711. doi: 10.1093/mnras/stw218.

Timothy M. Brown, David W. Latham, Mark E. Everett, and Gilbert A. Esquerdo. *Kepler*

Input Catalog: Photometric Calibration and Stellar Classification. *The Astronomical Journal*, 142(4):112, 2011. ISSN 0004-6256. doi: 10.1088/0004-6256/142/4/112.

Andrew J. Connolly, George Z. Angeli, Srinivasan Chandrasekharan, Charles F. Claver, Kem Cook, Zeljko Ivezic, R. Lynne Jones, K. Simon Krughoff, En-Hsin Peng, John Peterson, Catherine Petry, Andrew P. Rasmussen, Stephen T. Ridgway, Abhijit Saha, Glenn Sembroski, Jacob VanderPlas, and Peter Yoachim. An End-to-end Simulation Framework for the Large Synoptic Survey Telescope. *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 9150:915014, 2014. doi: 10.1117/12.2054953.

K. R. Covey, Ž. Ivezić, D. Schlegel, D. Finkbeiner, N. Padmanabhan, R. H. Lupton, M. A. Agüeros, J. J. Bochanski, S. L. Hawley, A. A. West, A. Seth, A. Kimball, S. M. Gogarten, M. Claire, D. Haggard, N. Kaib, D. P. Schneider, and B. Sesar. Stellar SEDs from 0.3 to 2.5 μm: Tracing the Stellar Locus and Searching for Color Outliers in the SDSS and 2MASS. *The Astronomical Journal*, 134(6):2398–2417, dec 2007. ISSN 0004-6256. doi: 10.1086/522052.

Francisco Delgado, Abhijit Saha, Srinivasan Chandrasekharan, Kem Cook, Catherine Petry, and Stephen Ridgway. The LSST Operations Simulator. In George Z. Angeli and Philippe Dierickx, editors, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 9150, page 915015, aug 2014. doi: 10.1117/12.2056898.

S. G. Djorgovski, A. J. Drake, A. A. Mahabal, M. J. Graham, C. Donalek, R. Williams, E. C. Beshore, S. M. Larson, J. Prieto, M. Catelan, E. Christensen, and R. H. McNaught. The Catalina Real-Time Transient Survey (CRTS). *arXiv preprint arXiv:* ..., page 6, feb 2011. URL http://arxiv.org/abs/1102.5004.

Amanda P. Doyle, B. Smalley, P. F. L. Maxted, D. R. Anderson, A. Collier Cameron, M. Gillon, C. Hellier, D. Pollacco, D. Queloz, A. H. M. J. Triaud, and R. G. West. Accurate spectroscopic parameters of WASP planet host stars. *Monthly Notices of the*

Royal Astronomical Society, 428(4):3164–3172, feb 2013. ISSN 0035-8711. doi: 10.1093/mnras/sts267.

B. Enoch, C. A. Haswell, A. J. Norton, A. Collier-Cameron, R. G. West, A. M. S. Smith, and N. R. Parley. Transit algorithm performance using real WASP data. *Astronomy & Astrophysics*, 548:A48, nov 2012. ISSN 0004-6361. doi: 10.1051/0004-6361/201219771.

H.C. Ferguson. Science White Paper for LSST Deep-Drilling Field Observations LSST Deep Drilling for Galaxies. *lsstcorp.org*, (410):1–9, 2011. URL https://lsstcorp.org/sites/default/files/WP/Ferguson-galaxies-02.pdf.

Debra A. Fischer, Geoffrey W. Marcy, and Julien F. P. Spronck. the Twenty-Five Year Lick Planet Search. *The Astrophysical Journal Supplement Series*, 210(1):5, jan 2014. ISSN 0067-0049. doi: 10.1088/0067-0049/210/1/5.

E. Gawiser, Jeff Newman, Hu Zhan, and David Ballantyne. Paper for LSST Deep-Drilling Field Observations Ultra-deep ugrizy Imaging to Reduce Main Survey Photo-z Systematics and to Probe Faint Galaxy Clustering,. *lsstweb.lsstcorp.org*, pages 1–11, 2011. URL https://lsstweb.lsstcorp.org/sites/default/files/WP/Gawiser-ultradeep-extragalactic-01.pdf.

L. Girardi, M. A. T. Groenewegen, E. Hatziminaoglou, and L. da Costa. Star counts in the Galaxy. Simulating from very deep to very shallow photometric surveys with the TRILEGAL code. *Astronomy & Astophysics*, 915(436):895–915, 2005. ISSN 0004-6361. doi: 10.1051/0004-6361:20042352.

Philip C. Gregory. Additional Keplerian Signals in the HARPS data for Gliese 667C from a Bayesian Re-analysis. *arXiv preprint arXiv:1212.4058*, (December):16, dec 2012. URL http://arxiv.org/abs/1212.4058.

J. Grindlay, S. Tang, R. Simcoe, S. Laycock, D. Mink, G. Champine, E. Los, and A. Doane. DASCH to Measure (and preserve) the Harvard Plates: Opening the 100-year Time Do-

main Astronomy Window. *In: Preserving Astronomy's Photographic Legacy: Current State and the Future of North American Astronomical Plates. ASP Conf. Series*, 410:101, 2009. URL http://adsabs.harvard.edu/cgi-bin/nph-data{_}query?bibcode=2009ASPC..410..101G{\&}link{_}type=ABSTRACT.

Jonathan Grindlay, Sumin Tang, Edward Los, and Mathieu Servillat. Opening the 100-Year Window for Time-Domain Astronomy. *Proceedings of the International Astronomical Union*, 7(S285):29–34, 2012. ISSN 1743-9213. doi: 10.1017/S1743921312000166.

G. R. Harp, Jon Richards, Seth Shostak, J. C. Tarter, Douglas A. Vakoch, and Chris Munson. Radio seti observations of the anomalous star kic 8462852. *The Astrophysical Journal*, 825(2):155, 2016.

J. D. Hartman, B. S. Gaudi, M. J. Holman, B. A. McLeod, K. Z. Stanek, J. A. Barranco, M. H. Pinsonneault, and J. S. Kalirai. Deep MMT Transit Survey of the Open Cluster M37. II. Variable Stars. *The Astrophysical Journal*, 675(2):1254–1277, mar 2008. ISSN 0004-637X. doi: 10.1086/527460.

Michael Hippke, Daniel Angerhausen, Michael B. Lund, Joshua Pepper, and Keivan G. Stassun. A statistical analysis of the accuracy of the digitized magnitudes of photometric plates on the timescale of decades with an application to the century-long light curve of kic 8462852. *The Astrophysical Journal*, 825(1):73, 2016.

S. T. Hodgkin, M. J. Irwin, P. C. Hewett, and S. J. Warren. The UKIRT wide field camera ZYJHK photometric system: calibration from 2MASS. *Monthly Notices of the Royal Astronomical Society*, 394(2):675–692, apr 2009. ISSN 00358711. doi: 10.1111/j.1365-2966.2008.14387.x.

Andrew W. Howard, Geoffrey W. Marcy, Stephen T. Bryson, Jon M. Jenkins, Jason F. Rowe, Natalie M. Batalha, William J. Borucki, David G. Koch, Edward W. Dunham, Thomas N. Gautier, Jeffrey Van Cleve, William D. Cochran, David W. Latham,

Jack J. Lissauer, Guillermo Torres, Timothy M. Brown, Ronald L. Gilliland, Lars A. Buchhave, Douglas A. Caldwell, Jørgen Christensen-Dalsgaard, David Ciardi, Francois Fressin, Michael R. Haas, Steve B. Howell, Hans Kjeldsen, Sara Seager, Leslie Rogers, Dimitar D. Sasselov, Jason H. Steffen, Gibor S. Basri, David Charbonneau, Jessie Christiansen, Bruce Clarke, Andrea Dupree, Daniel C. Fabrycky, Debra A. Fischer, Eric B. Ford, Jonathan J. Fortney, Jill Tarter, Forrest R. Girouard, Matthew J. Holman, John Asher Johnson, Todd C. Klaus, Pavel Machalek, Althea V. Moorhead, Robert C. Morehead, Darin Ragozzine, Peter Tenenbaum, Joseph D. Twicken, Samuel N. Quinn, Howard Isaacson, Avi Shporer, Philip W. Lucas, Lucianne M. Walkowicz, William F. Welsh, Alan Boss, Edna Devore, Alan Gould, Jeffrey C. Smith, Robert L. Morris, Andrej Prsa, Timothy D. Morton, Martin Still, Susan E. Thompson, Fergal Mullally, Michael Endl, and Phillip J. MacQueen. Planet Occurrence Within 0.25 Au of Solar-Type Stars From Kepler. *The Astrophysical Journal Supplement Series*, 201(2):15, aug 2012. ISSN 0067-0049. doi: 10.1088/0067-0049/201/2/15.

Z. Ivezic, J. A. Tyson, E. Acosta, R. Allsman, S. F. Anderson, J. Andrew, R. Angel, T. Axelrod, J. D. Barr, A. C. Becker, J. Becla, C. Beldica, R. D. Blandford, J. S. Bloom, K. Borne, W. N. Brandt, M. E. Brown, J. S. Bullock, D. L. Burke, S. Chandrasekharan, S. Chesley, C. F. Claver, A. Connolly, K. H. Cook, A. Cooray, K. R. Covey, C. Cribbs, R. Cutri, G. Daues, F. Delgado, H. Ferguson, E. Gawiser, J. C. Geary, P. Gee, M. Geha, R. R. Gibson, D. K. Gilmore, W. J. Gressler, C. Hogan, M. E. Huffer, S. H. Jacoby, B. Jain, J. G. Jernigan, R. L. Jones, M. Juric, S. M. Kahn, J. S. Kalirai, J. P. Kantor, R. Kessler, D. Kirkby, L. Knox, V. L. Krabbendam, S. Krughoff, S. Kulkarni, R. Lambert, D. Levine, M. Liang, K-T. Lim, R. H. Lupton, P. Marshall, S. Marshall, M. May, M. Miller, D. J. Mills, D. G. Monet, D. R. Neill, M. Nordby, P. O'Connor, J. Oliver, S. S. Olivier, K. Olsen, R. E. Owen, J. R. Peterson, C. E. Petry, F. Pierfederici, S. Pietrowicz, R. Pike, P. A. Pinto, R. Plante, V. Radeka, A. Rasmussen, S. T. Ridgway, W. Rosing, A. Saha, T. L. Schalk, R. H. Schindler, D. P. Schneider, G. Schumacher, J. Sebag,

L. G. Seppala, I. Shipsey, N. Silvestri, J. A. Smith, R. C. Smith, M. A. Strauss, C. W. Stubbs, D. Sweeney, A. Szalay, J. J. Thaler, D. Vanden Berk, L. Walkowicz, M. Warner, B. Willman, D. Wittman, S. C. Wolff, W. M. Wood-Vasey, P. Yoachim, H. Zhan, and for the LSST Collaboration. LSST: from Science Drivers to Reference Design and Anticipated Data Products. page 34, may 2008. URL http://arxiv.org/abs/0805.2366.

Savannah R. Jacklin, Michael B. Lund, Joshua Pepper, and Keivan G. Stassun. Transiting Planets with LSST II. Period Detection of Planets Orbiting 1 Solar Mass Hosts. *The Astronomical Journal*, 150(1):34, 2015. ISSN 1538-3881. doi: 10.1088/0004-6256/150/1/34.

T. Jackson, Z. Ivezic, and G. R. Knapp. The Galactic distribution of asymptotic giant branch stars. *Monthly Notices of the Royal Astronomical Society*, 337(2):749–767, dec 2002. ISSN 0035-8711. doi: 10.1046/j.1365-8711.2002.05980.x.

C. Jordi and J.M. Carrasco. Photometry with GAIA. *Astronomical Society of the Pacific Conference Series*, 364:215, 2006. URL http://www.am.ub.edu/{~}carrasco/documentsgaia/215.pdf.

Mario Jurić, Željko Ivezić, Alyson Brooks, Robert H. Lupton, David Schlegel, Douglas Finkbeiner, Nikhil Padmanabhan, Nicholas Bond, Branimir Sesar, Constance M. Rockosi, Gillian R. Knapp, James E. Gunn, Takahiro Sumi, Donald P. Schneider, J. C. Barentine, Howard J. Brewington, J. Brinkmann, Masataka Fukugita, Michael Harvanek, S. J. Kleinman, Jurek Krzesinski, Dan Long, Eric H. Neilsen, Jr., Atsuko Nitta, Stephanie A. Snedden, and Donald G. York. The Milky Way Tomography with SDSS. I. Stellar Number Density Distribution. *The Astrophysical Journal*, 673(2):864–914, feb 2008. ISSN 0004-637X. doi: 10.1086/523619.

S. Kafka. Observations from the AAVSO International Database, 2015. URL http://www.aavso.org.

Nicholas Kaiser. Pan-STARRS: a wide-field optical survey telescope array. In Jacobus M. Oschmann, Jr., editor, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 5489, pages 11–22, sep 2004. doi: 10.1117/12.552472. URL http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1317240.

M. G. Kendall. A New Measure of Rank Correlation. *Biometrika*, 30(1/2):81, jun 1938. ISSN 00063444. doi: 10.2307/2332226.

Maciej Konacki, Guillermo Torres, Saurabh Jha, and Dimitar D. Sasselov. An extrasolar planet that transits the disk of its parent star. *Nature*, 421(6922):507–509, jan 2003. ISSN 0028-0836. doi: 10.1038/nature01379.

G. Kovacs, S. Zucker, and T. Mazeh. A box-fitting algorithm in the search for periodic transits. *Astronomy and Astrophysics*, 391(1):369–377, aug 2002. ISSN 0004-6361. doi: 10.1051/0004-6361:20020802.

S. Laycock, S. Tang, J. Grindlay, E. Los, R. Simcoe, and D. Mink. Digital Access To a Sky Century At Harvard: Initial Photometry and Astrometry. *The Astronomical Journal*, 140 (4):1062–1077, 2010. ISSN 0004-6256. doi: 10.1088/0004-6256/140/4/1062.

N. R. Lomb. Least-squares frequency analysis of unequally spaced data. *Astrophysics and Space Science*, 39(2):447–462, 1976. ISSN 0004640X. doi: 10.1007/BF00648343.

LSST Science Collaboration, P.A. Abell, J. Allison, S.F. Anderson, J.R. Andrew, J.R.P. Angel, L. Armus, D. Arnett, S.J. Asztalos, T.S. Axelrod, and Et al. *LSST Science Book*. LSST Science Collaboration, Tucson, AZ, version 2. edition, 2009. URL http://www.lsst.org/lsst/scibook.

Michael B. Lund. Astrology in the Era of Exoplanets. II(April), 2016. URL http://arxiv.org/abs/1603.09496.

Michael B. Lund, Joshua Pepper, and Keivan G. Stassun. Transiting Planets With LSST. I. Potential for LSST Exoplanet Detection. *The Astronomical Journal*, 149(1):16, dec 2015. ISSN 1538-3881. doi: 10.1088/0004-6256/149/1/16.

Michael B. Lund, Joshua Pepper, Keivan G. Stassun, Michael Hippke, and Daniel Angerhausen. The Stability of F-star Brightness on Century Timescales. page 14, 2016a. URL http://arxiv.org/abs/1605.02760.

Michael B. Lund, Robert J. Siverd, Joshua A. Pepper, and Keivan G. Stassun. Metrics for Optimization of Large Synoptic Survey Telescope Observations of Stellar Variables and Transients. *Publications of the Astronomical Society of the Pacific*, 128(960):025002, 2016b. ISSN 0004-6280. doi: 10.1088/1538-3873/128/960/025002.

Massimo Marengo, Alan Hulsebus, and Sarah Willis. KIC 8462852: the Infrared Flux. *The Astrophysical Journal*, 814(1):L15, 2015. ISSN 2041-8213. doi: 10.1088/2041-8205/814/1/L15.

M. Mayor and D. Queloz. A Jupiter-mass companion to a solar-type star. *Nature*, 378: 355–359, 1995. ISSN 0028-0836. doi: 10.1038/378355a0.

Francis T. O'Donovan and David Charbonneau. TrES Exoplanets and False Positives: Finding the Needle in the Haystack. *arXiv preprint arXiv:0705.1795*, 366(Pollack 1984): 6, may 2007. URL http://arxiv.org/abs/0705.1795.

Shannon G. Patel, Steven S. Vogt, Geoffrey W. Marcy, John A. Johnson, Debra A. Fischer, Jason T. Wright, and R. Paul Butler. Fourteen New Companions from the Keck and Lick Radial Velocity Survey Including Five Brown Dwarf Candidates. *The Astrophysical Journal*, 665(1):744–753, aug 2007. ISSN 0004-637X. doi: 10.1086/519066.

Karl Pearson. X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to

have arisen from random sampling. *Philosophical Magazine Series 5*, 50(302):157–175, jul 1900. ISSN 1941-5982. doi: 10.1080/14786440009463897.

Joshua Pepper, Richard W. Pogge, D. L. DePoy, J. L. Marshall, K. Z. Stanek, Amelia M. Stutz, Shawn Poindexter, Robert Siverd, Thomas P. O'Brien, Mark Trueblood, and Patricia Trueblood. The Kilodegree Extremely Little Telescope (KELT): A Small Robotic Telescope for LargeArea Synoptic Surveys. *Publications of the Astronomical Society of the Pacific*, 119(858):923–935, aug 2007. ISSN 0004-6280. doi: 10.1086/521836.

Erik A. Petigura, Andrew W. Howard, and Geoffrey W. Marcy. Prevalence of Earth-size planets orbiting Sun-like stars. *Proceedings of the National Academy of Sciences of the United States of America*, 2013(September):1–6, nov 2013. ISSN 1091-6490. doi: 10.1073/pnas.1319909110.

Radosaw Poleski, Peter R. McCullough, Jeff A. Valenti, Christopher J. Burke, Pavel Machalek, and Kenneth Janes. the XO Planetary Survey Project: Astrophysical False Positives. *The Astrophysical Journal Supplement Series*, 189(1):134–141, jul 2010. ISSN 0067-0049. doi: 10.1088/0067-0049/189/1/134.

Arne Rau, Shrinivas R. Kulkarni, Nicholas M. Law, Joshua S. Bloom, David Ciardi, George S. Djorgovski, Derek B. Fox, Avishay Gal-Yam, Carl C. Grillmair, Mansi M. Kasliwal, Peter E. Nugent, Eran O. Ofek, Robert M. Quimby, William T. Reach, Michael Shara, Lars Bildsten, S. Bradley Cenko, Andrew J. Drake, Alexei V. Filippenko, David J. Helfand, George Helou, D. Andrew Howell, Dovi Poznanski, and Mark Sullivan. Exploring the Optical Transient Sky with the Palomar Transient Factory. *Publications of the Astronomical Society of the Pacific*, 121(886):1334–1351, dec 2009. ISSN 0004-6280. doi: 10.1086/605911.

D. H. Roberts, J. Lehar, and J. W. Dreher. Time Series Analysis with Clean - Part One -

Derivation of a Spectrum. *The Astronomical Journal*, 93:968, 1987. ISSN 00046256. doi: 10.1086/114383.

Joseph E. Rodriguez, Keivan G. Stassun, Michael B. Lund, Robert J. Siverd, Joshua Pepper, Sumin Tang, Stella Kafka, Scott Gaudi, Kyle E. Conroy, Thomas G. Beatty, and Daniel J. Stevens. An Extreme Analogue of ϵ Aurigae: An M-giant Eclipsed Every 69 Years by a Large Opaque Disk Surrounding a Small Hot Source. *The Astronomical Journal*, 151 (5):123, 2016.

J. D. Scargle. Studies in astronomical time series analysis. II - Statistical aspects of spectral analysis of unevenly spaced data. *The Astrophysical Journal*, 263:835, dec 1982. ISSN 0004-637X. doi: 10.1086/160554.

Bradley E. Schaefer. Kic 8462852 faded at an average rate of 0.164 0.013 magnitudes per century from 1890 to 1989. *The Astrophysical Journal Letters*, 822(2):L34, 2016.

Marlin Schuetz, Douglas A. Vakoch, Seth Shostak, and Jon Richards. Optical seti observations of the anomalous star kic 8462852. *The Astrophysical Journal Letters*, 825(1):L5, 2016.

R.J. Simcoe, J.E. Grindlay, E.J. Los, A. Doane, S.G. Laycock, D.J. Mink, G. Champine, and A. Sliski. An ultrahigh-speed digitizer for the Harvard College Observatory astronomical plates. *Proc. SPIE*, 6312:631217–631217–12, 2006. doi: 10.1117/12.681365. URL http://dx.doi.org/10.1117/12.681365.

P. Szkody, K.S. Long, R. DiStefano, A. Henden, J. Kalirai, V. Kashyap, M. Kasliwal, J. A. Smith, and K. Stassun. Science White Paper for LSST Deep-Drilling Field Observations High Cadence Observations of the Magellanic Clouds and Select Galactic Cluster Fields. *lsstcorp.org*, pages 1–10, 2011. URL https://lsstcorp.org/sites/default/files/WP/Szokody-MCs-gal-clusters-01.pdf.

Sumin Tang, Jonathan Grindlay, Edward Los, and Silas Laycock. Dasch Discovery of Large Amplitude 10-100 Year Variability in K Giants. *The Astrophysical Journal*, 710 (1):L77–L81, 2010. ISSN 2041-8205. doi: 10.1088/2041-8205/710/1/L77.

Sumin Tang, Jonathan Grindlay, Edward Los, and Mathieu Servillat. Improved Photometry for the DASCH Pipeline. *Publications of the Astronomical Society of the Pacific*, 125 (929):857–865, 2013a. ISSN 00046280. doi: 10.1086/671760.

Sumin Tang, Dimitar Sasselov, Jonathan Grindlay, Edward Los, and Mathieu Servillat. 100-year DASCH Light Curves of Kepler Planet-Candidate Host Stars. *Publications of the Astronomical Society of the Pacific*, 125(929):793, 2013b. ISSN 00046280. doi: 10.1086/671759.

M. A. Thompson, P. Scicluna, F. Kemper, J. E. Geach, M. M. Dunham, O. Morata, S. Ertel, P. T. P. Ho, J. Dempsey, I. Coulson, G. Petitpas, and L. E. Kristensen. Constraints on the circumstellar dust around KIC 8462852. *Monthly Notices of the Royal Astronomical Society*, 458(1):L39, 2016. ISSN 1745-3925. doi: 10.1093/mnrasl/slw008.

C. G. Tinney, R. Paul Butler, Geoffrey W. Marcy, Hugh R. A. Jones, Alan J. Penny, Steven S. Vogt, Kevin Apps, and Gregory W. Henry. First Results from the AngloAustralian Planet Search: A Brown Dwarf Candidate and a 51 Peglike Planet. *The Astrophysical Journal*, 551(1):507–511, apr 2001. ISSN 0004-637X. doi: 10.1086/320097.

B. L. Welch. The Generalization Of Student's' Problem When Several Different Population Varlances are Involved. *Biometrika*, 34(1-2):28–35, 1947. ISSN 0006-3444. doi: 10.1093/biomet/34.1-2.28.

A. Wolszczan and D. A. Frail. A planetary system around the millisecond pulsar PSR1257 + 12. *Nature*, 355(6356):145–147, jan 1992. ISSN 0028-0836. doi: 10.1038/355145a0.

Candace O. Wright, Michael P. Egan, Kathleen E. Kraemer, and Stephan D. Price. The

Tycho-2 Spectral Type Catalog. *The Astronomical Journal*, 125(1):359, 2003. ISSN 00046256. doi: 10.1086/345511.

Jason T. Wright, Kimberly M. S. Cartier, Ming Zhao, Daniel Jontof-Hutter, and Eric B. Ford. The $\hat{\mathrm{G}}$ search for extraterrestrial civilizations with large energy supplies. iv. the signatures and information content of transiting megastructures. *The Astrophysical Journal*, 816(1):17, 2015. ISSN 1538-4357. doi: 10.3847/0004-637X/816/1/17.

Andrew N. Youdin. The Exoplanet Census: a General Method Applied To Kepler. *The Astrophysical Journal*, 742(1):38, nov 2011. ISSN 0004-637X. doi: 10.1088/0004-637X/742/1/38.

M. Zechmeister, M. Kürster, M. Endl, G. Lo Curto, H. Hartman, H. Nilsson, T. Henning, A. P. Hatzes, and W. D. Cochran. The planet search programme at the ESO CES and HARPS. *Astronomy & Astrophysics*, 552(December):A78, apr 2013. ISSN 0004-6361. doi: 10.1051/0004-6361/201116551.

Appendix A

Stars with Significant Slopes

In the analysis in this paper, we examined over 600 stars, and we fit over 1200 lines to the light curves after the data is split at 1970. If the non-zero slopes of these lines are purely a result of statistical noise in flat light curves, the chance of having even one slope that is significant by more than 5 σ is extremely small. Instead, we find 11 light curves that exceed this threshold. These stars may simply be indicative of further systematics that we have not addressed, but we present them here in Figures A.1 and A.2 in the case that these may also represent astrophysical behavior.

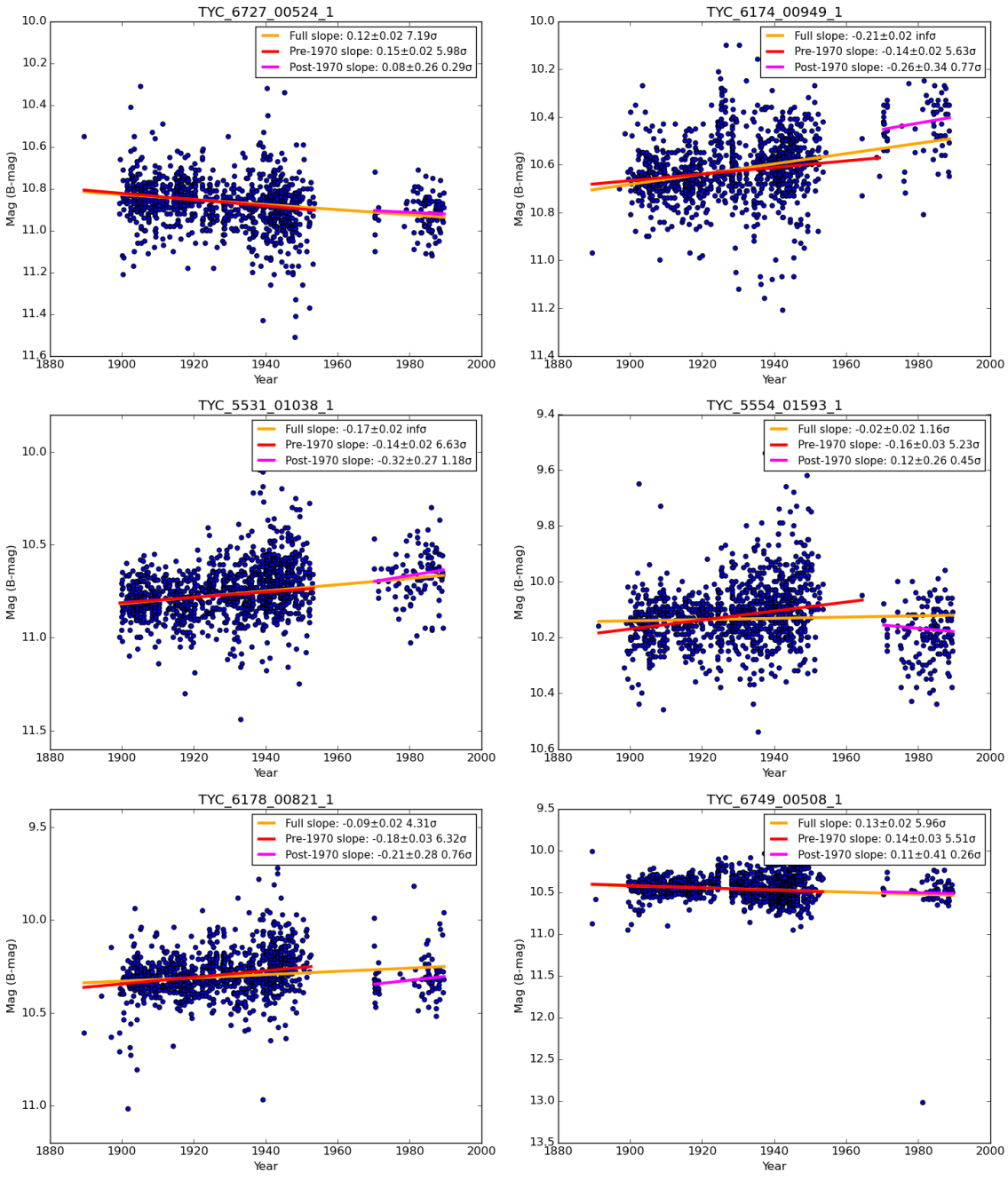

Figure A.1: Stars showing slopes greater than 5 σ from 0 either pre-1970 or post-1970.

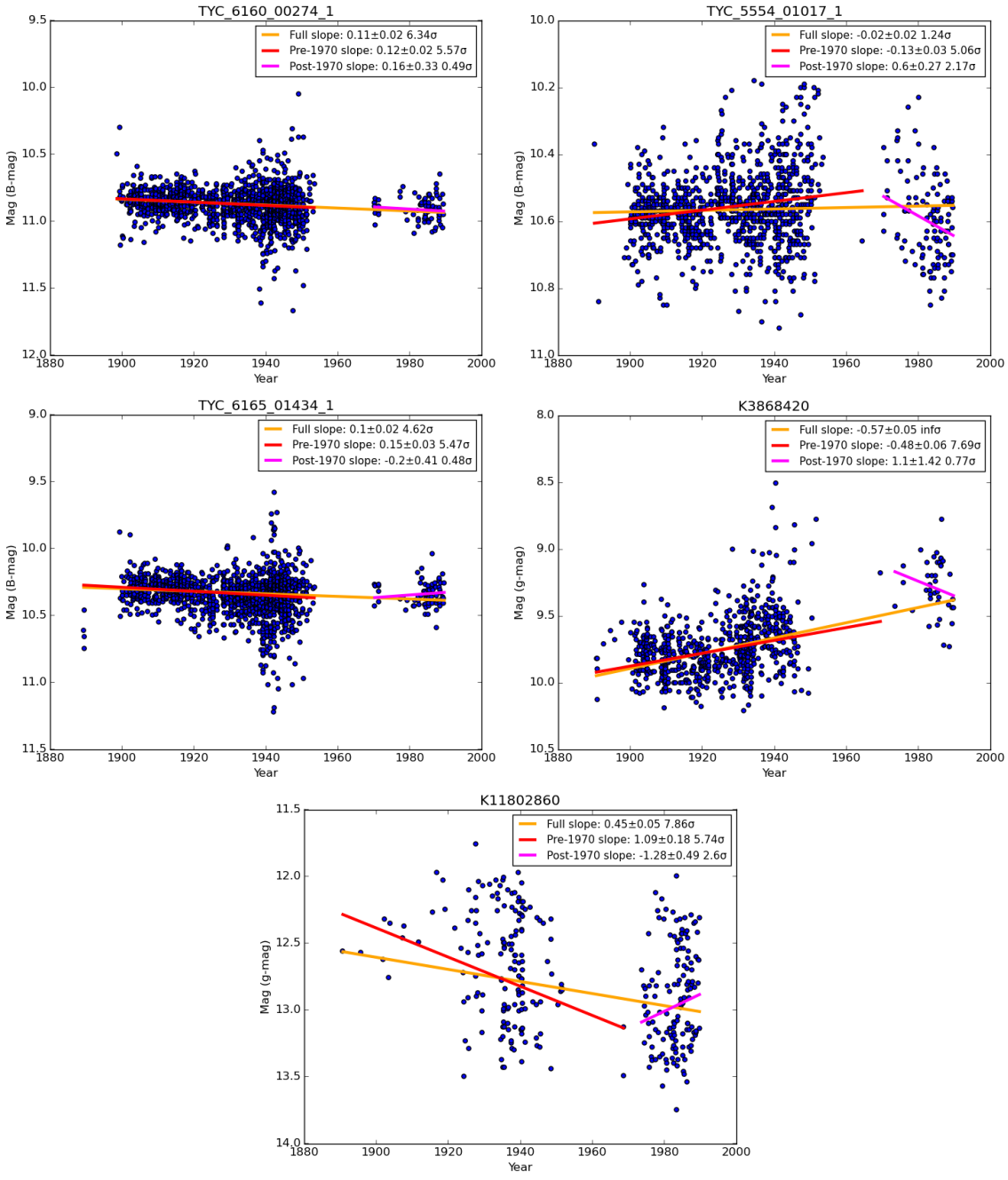

Figure A.2: Stars showing slopes greater than 5 σ from 0 either pre-1970 or post-1970.