
Wesleyan University

Removing Contaminating Field Stars Around Bright *K2* Targets

by

Katharine Hesse
Class of 2020

A thesis submitted to the
faculty of Wesleyan University
in partial fulfillment of the requirements for the
Degree of Master of Arts

Middletown, Connecticut

May, 2020

“It’s only through honesty and courage that science can work at all.”

—PHILIP PULLMAN

Acknowledgements

To Seth, thank you for always wholly supporting my progress, even and especially when it felt so small a step. You taught me the results often do not match the goals we set and the process will always be more complex than expected, and that this is normal and part of science. Anything worth working on takes significant time and effort.

To Ismael, thank you for serving as my *K2* expert and always having the patience to explain questions about your past work and python coding as a whole. Whether asking direct questions or just having someone to talk through my process with, I could not have put this project together without your support.

To Jessica, thank you for helping me get through my first year of graduate school. Going back to school was a rougher transition than I was prepared for and you understood that and were always there with advice and a calm, practical outlook.

To Justin, thank you for always being available to bounce questions and ideas off of, both for thesis and coursework. I appreciated the many long nights spent working in the office, and appreciated still having the spirit of that crunch time through texts.

To Tedi, thank you for being understanding about my constant levels of stress and having my back when all the work felt so completely overwhelming. You kept me sane through the entirety of my thesis, graduate applications, and degree and I could not have asked for a better partner to be with through this. You gave me a newfound appreciation for my own work, and always reminded me I was accomplishing so much even though it never felt like enough. You encouraged me to be proud of myself and my work and always believed in me.

To Anjel, thank you for helping me with last minute thesis edits and sitting through and reading the whole thing, no matter how late it was. Throughout these last few weeks you have been an invaluable listener and source of advice.

To Sara, Chelsea, Jenn, Katie, and Danielle, thank you for cheering me on and supporting my progress. Every comment or thumbs up or conversation about my project gave me strength to keep going.

To my cat, Lyra, thank you for providing a good cuddle when times were rough. Especially in the final push, having you sitting on the bed near my desk and being able to take quick break for a kiss and a belly rub was the perfect rejuvenation.

To my family, thank you for always checking in on me, even and especially when I got caught up in work and was not the most available. Your constant support and pride in me and my work meant the world and knowing I can always count on that will keep me going in the next steps forward.

To the Wesleyan Astronomy department as a whole, thank you for being so inclusive and welcoming. Between colloquium, classes, seminars, Astro beers, and even just the daily greetings and farewells in the hallways, I always felt welcome. Our time in person ended too soon but the sense of community has still been so present through technology and it has helped me through to the end. I could not have asked for a better, more supportive place to complete my Master's.

Contents

Acknowledgements	ii
1 Introduction	1
1.1 The Search So Far	1
1.1.1 Radial Velocity Detection Method	3
1.1.2 Transit Detection Method	4
1.1.3 Challenges for Transit Method	9
1.1.4 <i>Kepler</i> Mission and <i>K2</i> Rebirth	11
1.1.5 <i>TESS</i>	13
1.1.6 Future Missions	15
1.2 Bright Stars	16
1.2.1 Halophot	17
1.3 PSF Modeling	19
1.3.1 Double Gaussian	20
1.3.2 Halophot: Deathstar	21
1.3.3 PSF Model Creation	22
1.4 Goals for PSF Model	27
1.4.1 PSF Variance	27
1.4.2 Exoplanets Around Bright Stars	28
1.4.3 Exoplanets Around Background Stars	28

2	Data: <i>Kepler</i> vs. <i>K2</i>	29
2.1	The Instrument	29
2.2	Observations and Data Processing	34
2.3	Field of View	38
2.4	Motion	43
2.4.1	<i>K2</i> Thruster Firing	44
2.5	Summary of <i>K2</i> vs. <i>Kepler</i> and the Effect on PSF	46
3	Methodology	47
3.1	PRF vs. PSF	47
3.2	Contaminating Stars	49
3.3	PSF Modeling	55
3.3.1	PSF Orientation	58
3.4	Summary of Approach to <i>K2</i> PSF Creation	58
4	Results	60
4.1	Single Star Observations By Channel and Campaign	61
4.2	PSF Stacks by Channel	65
4.3	Discussion of Initial Results	72
4.3.1	Campaign 1	72
4.3.2	Campaign 7	73
4.3.3	Campaign 13	74
4.4	Evaluating the PSF Model	77
4.4.1	PSF vs. PRF	78
4.4.2	PSF vs. Gaussian	79
4.4.3	Comparing PSF between Campaigns by Channel	79
4.5	Future Steps	88

5 PSF Model Applications	89
5.1 Application to Halophot	89
5.2 Transits Around Pre-processed Bright Stars	90
6 Conclusion	95
6.1 Future Steps for PSF Modeling	95
6.2 Applying PSF Models Bright Stars	97
6.3 Using Background Stars	97
6.4 Bright Stars in TESS	98
Bibliography	99

Chapter 1

Introduction

1.1 The Search So Far

The existence of habitable planets has long been considered, although the research in the field of exoplanets has only recently picked up steam. Once Copernicus had shared the helio-centric model in which planets orbit the Earth, other philosophers began suspecting that there may be planets orbiting other stars. In the 16th century, the philosopher Giordano Bruno extended the Copernican model by proposing that infinite stars like the Sun exist, each hosting their own planets (Maor and Stewart 2017). Given the existence of the solar system, it is only logical to assume there are many stars with similar solar systems to our own. What used to be the content of pure imagination became a reality that could be pursued in 1952 when Otto Sturwe wrote a paper detailing the process for radial velocity and transit detection methods for exoplanets (Deeg and Alonso 2018). These methods have become fundamental to the search for extra-solar planets and numerous examples have been found. Figure 1.1 is a histogram displaying the frequency of detections with each method over time. Radial velocity has been a consistently productive method for detections, but the transit method has been increasing in the amount of detections since 2010. This is around when *Kepler* started returning data.

1. INTRODUCTION

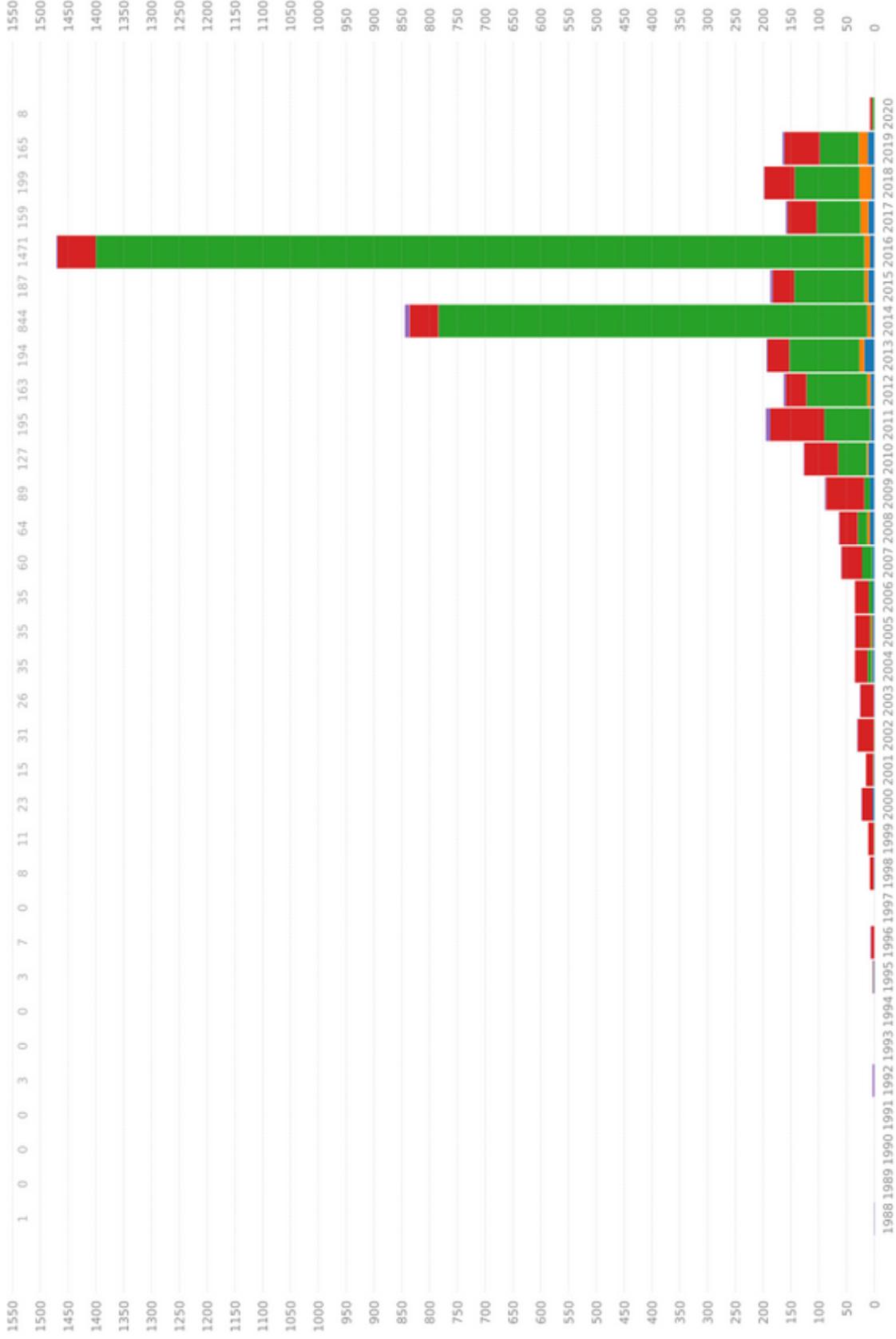


Figure 1.1: Histogram showing exoplanet detection by method. Red indicates radial velocity, green indicates transit, blue indicates direct imaging, orange indicates microlensing, and purple indicates timing. Plot credit: https://en.wikipedia.org/wiki/Methods_of_detecting_exoplanets

1.1.1 Radial Velocity Detection Method

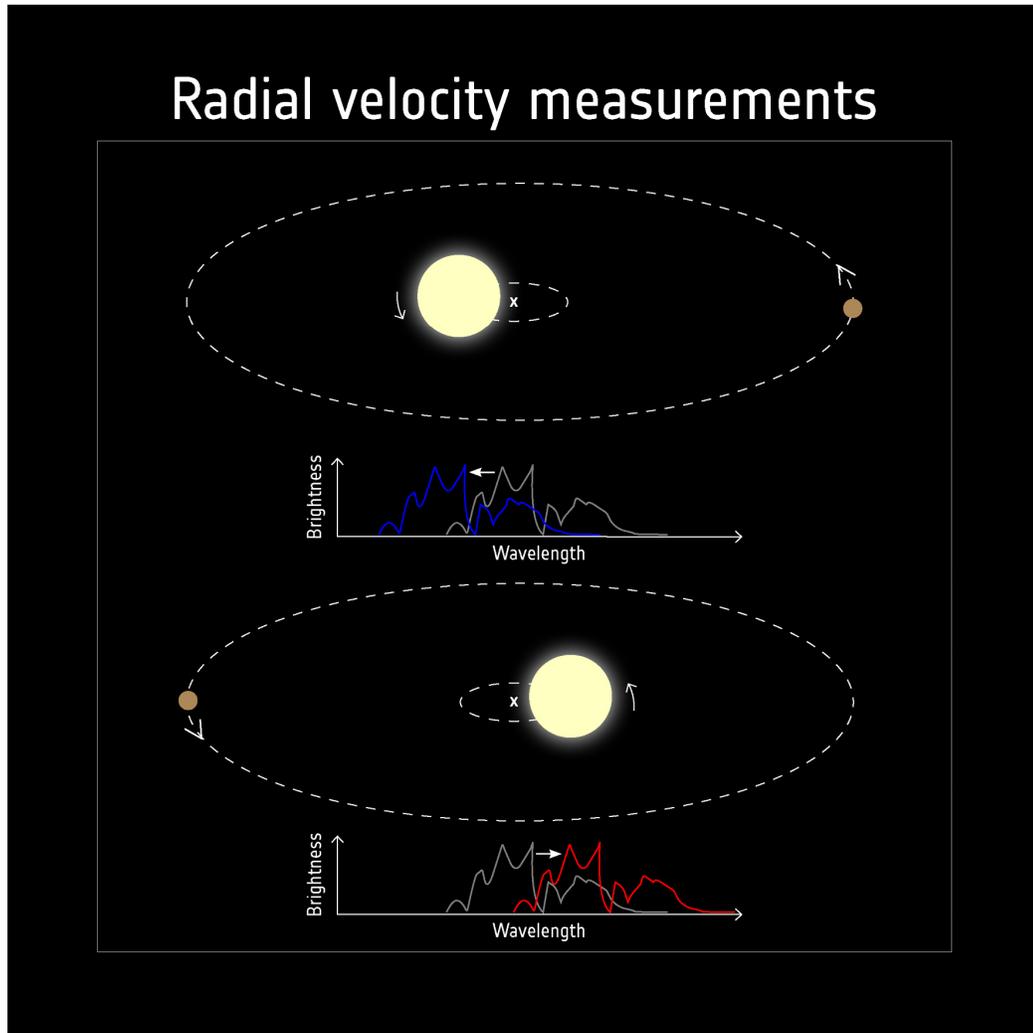


Figure 1.2: Diagram showing how the radial velocity detection method works. When the unseen planet moves away from us, the star moves closer and its spectrum is blue shifted. When the unseen planet moves closer to us, the star moves away and its spectrum is red shifted. Image credit: https://www.esa.int/ESA_Multimedia/Images/2019/02/Detecting_exoplanets_with_radial_velocity

Radial velocity detection involves the fact that when orbited by a planet, a star will experience a slight gravitational pull from the planet. This gravitational pull causes the star to move back and forth slightly, its center of mass moving in a very small ellipse, and the presence of the (typically large) planet causes the center

of gravity for the system to not be at the center of the star. When observed, this slight motion corresponds to shifts in the spectrum of the star due to the Doppler Effect. Both the planet and the star orbit the same center of gravity and orbit on opposite sides of the center of gravity. When the planet is moving away from the viewer, the star moves towards the viewer, causing the light from the star to compress and the spectrum to be blue-shifted. When the planet is moving towards the viewer, the star moves away from the viewer, causing the light from the star to stretch and the spectrum to be red-shifted. Even though the planet cannot be directly seen, its presence and mass can be determined based on its gravitational effects on the star's spectrum.

However, since this shift is strongest in cases where the planet is massive and closer to the star, these types of exoplanets have been over represented using this method. The radial velocity method is also unable to determine the radius of the planet. Other methods have since been developed, such as the transit detection method that offer more information about the extra-solar planets around other stars.

1.1.2 Transit Detection Method

Since around 2010, transit photometry has become the most prolific method of exoplanet detection. This method can determine a wide range of characteristics for an exoplanet, such as radius, orbital parameters like semi-major axis and period, and even atmosphere characterization, which the radial velocity method cannot measure.

The transit method is used to detect exoplanets by measuring variation and reduction in stellar flux as the planet passes in front of the host star. This process

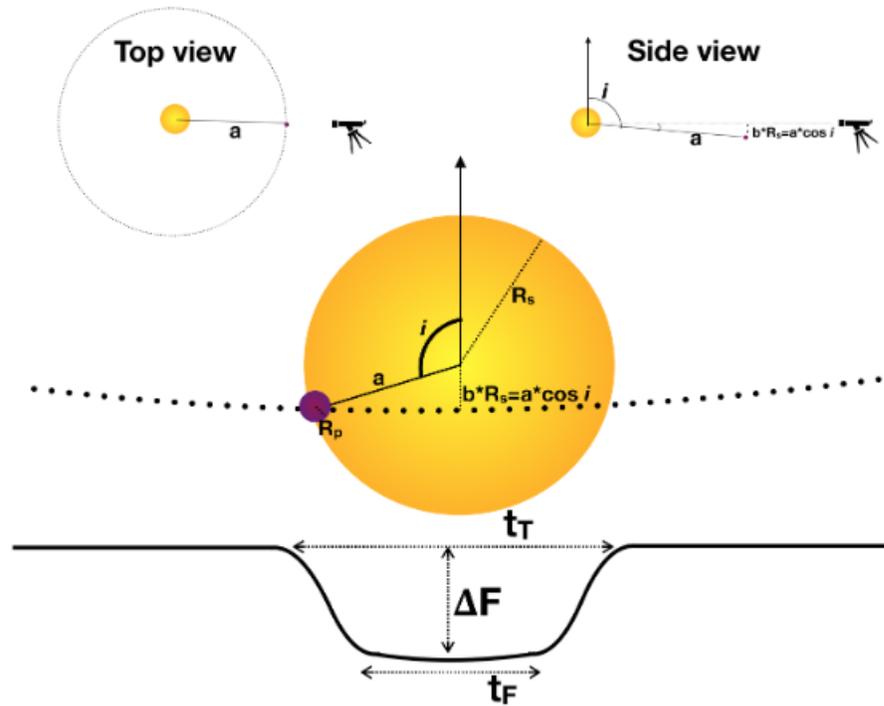


Figure 1.3: Diagram labeling characteristics describing an exoplanet’s transit across its host star, from Deeg and Alonso (2018) .

is shown in Figure 1.3, where R_p is the radius of the planet, R_s is the radius of the star, i is the inclination of the planet’s orbit, Δf is the change in flux, a is the semi-major axis of the planet’s orbit, and t_T is the duration of the transit.

To determine the reduction in flux due to the transit, a ratio is made between the square of the radius of the planet to the square of the radius of the star. Simply put, this is the calculation of how much of the star’s cross-section is being covered by the cross-section of the planet as it transits the star. The radius of the star can be estimated based off its spectral type, and the change in flux (Δf) can be measured by taking a flux time series for the star. Thus, setting the change in flux equal to the ratio of the stellar radius to the planet radius allows the planet’s radius to be calculated with two simple properties, change in flux and

stellar radius. This concept is put into an equation in the following form:

$$\Delta F \approx \left(\frac{R_p}{R_s}\right)^2 = k^2 \quad (1.1)$$

where the change in stellar flux is equal to the ratio of the planet's radius to star's radius squared, while k is referred to as the transit ratio (Deeg and Alonso 2018). This transit ratio, simply the ratio between the planet radius and stellar radius, can be used to calculate additional planet parameters.

More information can be calculated about the planet's orbit, using the transit ratio along with another transit feature called the time of totality (t_R). The time of totality is the duration of the lowest level of flux during a transit. In other words, it is how long the planet is fully blocking the light from the star. There are transitional periods, when the planet first begins covering the star but is not fully in front of it at the starting edge, and similarly when the planet is at the trailing edge, where only part of the planet's cross-section is overlaid and blocking light from the sun.

The time of totality, along with the transit features previously described above, can be used to calculate the impact parameter b . The impact parameter is the smallest projected distance between the orbiting planet and the center of the stellar disk. In Figure 1.4, a planet is shown about to transit in front of a star. The impact parameter is a measure of how close the planet gets to the center of the star, marked with a dot in this image. An impact parameter of $b = 0$ would indicate a transit in which the planet crosses the center of the stellar disk. An impact of parameter of $b = 1$ would indicate a planet crossing right at the edge of the stellar disk. The transit for a planet would be the longest at $b = 0$, since the planet is crossing through the largest area across the cross-section of the star.

Any impact parameters higher than 0 would correspond to a shorter period, since the planet would be crossing through a shorter length along the cross-section of the star.

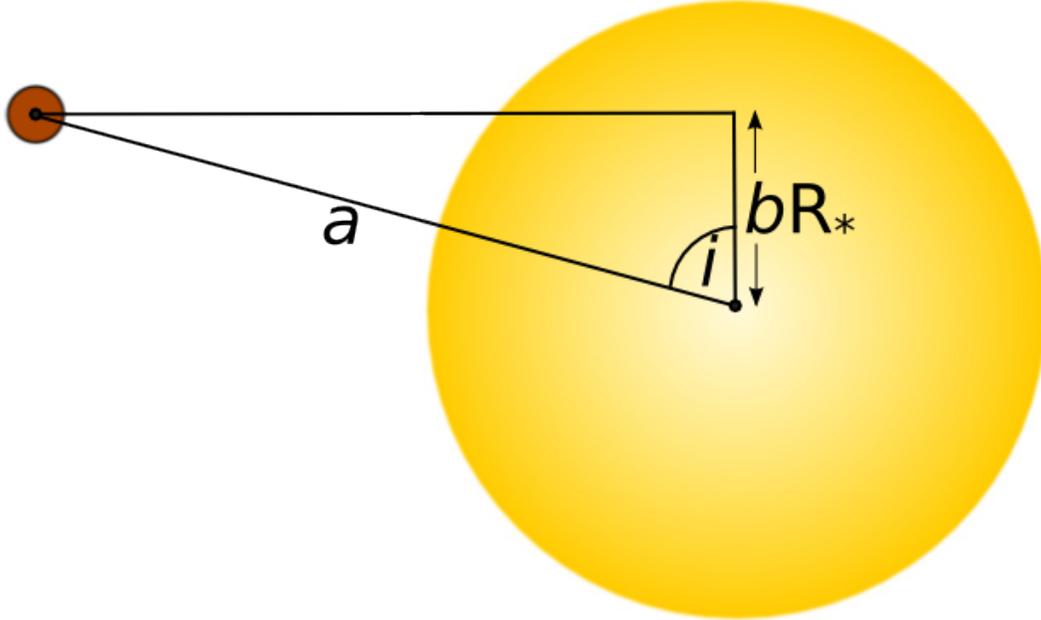


Figure 1.4: Diagram showing the definition of the impact parameter, b . Image credit: <https://www.paulanthonywilson.com/exoplanets/exoplanet-detection-techniques/the-exoplanet-transit-method/>.

The impact parameter can be calculated using the following equation:

$$b = \frac{a}{R_s} \cos i = \left(\frac{(1 - k)^2 - [\sin^2(t_f \pi / P) / \sin^2(t_T \pi / P)] (1 + k)^2}{\cos^2(t_f \pi / P) / \cos^2(t_T \pi / P)} \right)^{1/2} \quad (1.2)$$

where a is the semi-major orbital axis, i is the orbital inclination, P is the orbital period, t_F is the time of totality for the transit as previously described, and t_T is the total time of transit (Deeg and Alonso 2018). The equation can be used to calculate the impact parameter, as well as the ratio of the semi-major orbital axis to the radius of the star, referred to as the scale of the system. This ratio can be

used to determine the average stellar density of a host star, assuming a circular orbit between the planet and host star and that the star is much more massive than the planet, using the following equation:

$$\rho_s = \frac{3\pi}{GP^2} * \left(\frac{a}{R_s}\right)^3 \quad (1.3)$$

where ρ_s is the stellar density and G is the gravitational constant (Deeg and Alonso 2018). Knowing the density of a star indicates a lot about what type of star is being observed, as well as what stage of its life cycle the star is in. Planets around stars similar to the Sun or around other low-mass main sequence stars are ideal candidates for follow up, as they have long main sequence life-times and have a longer time-frame for some form of life to develop than higher-mass main sequence stars or any evolved stars. Evolved stars in particular would be poor candidates for habitable planets, since they are much more variable than main sequence stars.

With the transit method, much more information can be determined about extra-solar planets than is possible with the radial velocity method. The radial velocity method cannot determine the radius of a planet, nor additional characterization features like the presence or potential components of the atmosphere.

The transit method is also capable of identifying smaller planets than the radial velocity method. For the radial velocity method to accurately find a star, the companion planet must be close and large enough to significantly influence the center of mass in the system. The transit method, on the other hand, only requires the planet to be large enough in proportion to the star to significantly reduce the flux of the star for a transit to be observed. While this method is still biased towards closer and larger planets, as they would cover more of the star and

lead to a larger reduction in stellar flux, the range of sizes that the transit method can detect is much larger than the radial velocity method.

1.1.3 Challenges for Transit Method

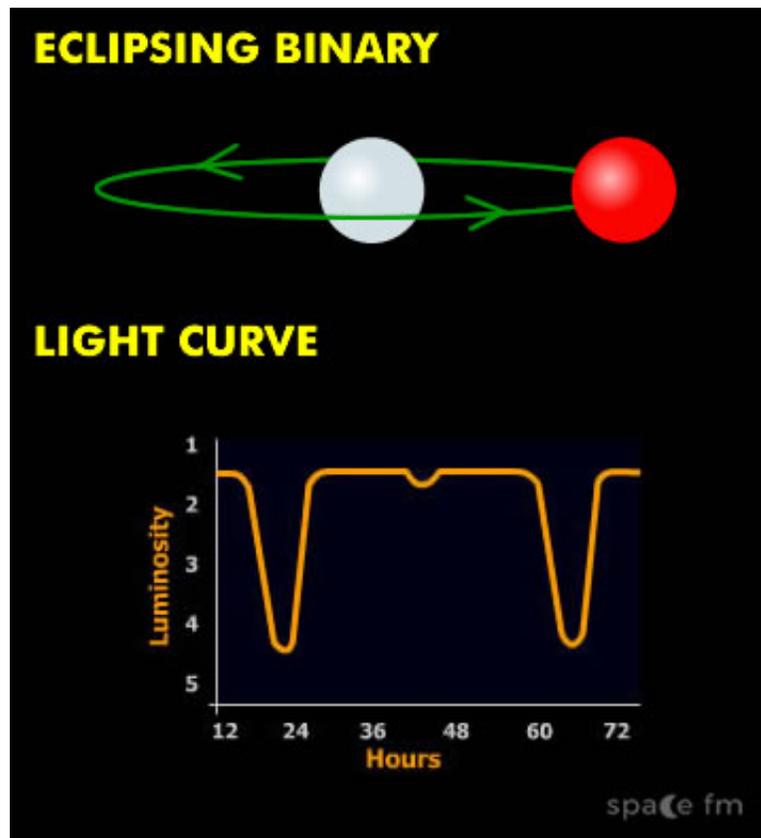


Figure 1.5: Diagram showing the light curve of an eclipsing binary. It is notably different to a planet transit since the transits are more “V”-shaped and there is a small, secondary reduction in flux. Image credit: <https://www.space.fm/astromy/starsgalaxies/eclipsingbinarylightcurve.html> .

The transit method has significant advantages over the radial velocity method, but there are also drawbacks. Apparent transits in a light curve can be due to sources other than an exoplanet passing in front of the star, resulting in false positives.

Eclipsing binaries can cause a transit-like reduction in flux when one star crosses in front of the other. When a star crosses in front of the other, it is still blocking light from the star in the background. However, this reduction in flux is often more narrow, resulting in more of “V” shape than a transiting planet. There is also a large difference between the odd and even transits, since one star is typically brighter than the other. Both of these features can be seen in Figure 1.5.

False positives can arise through other unique situations involving binaries as well. One scenario could be an eclipsing binary in the background of the image star, potentially in a triple system. Another scenario could be a planet around the other star in the binary (Santerne et al. 2013). These scenarios are illustrated in Figure 1.6, along with a typical eclipsing binary.

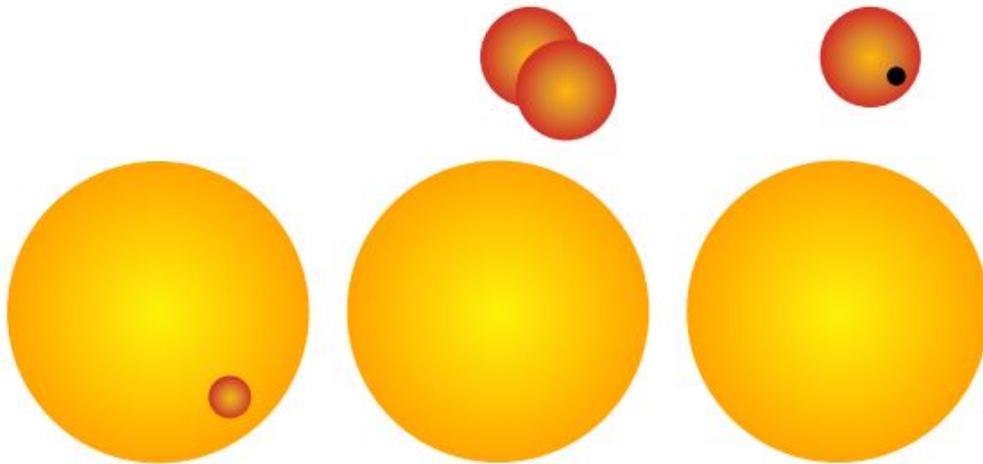


Figure 1.6: Diagram showing possible false positive scenarios. The first image is an eclipsing binary. The second is a background eclipsing binary or a binary in an eclipsing system. The third is a transiting planet around a background star or a companion star. Image credit: Santerne et al. (2013) .

The best way to check if these scenarios above are the case for a suspected transiting planet is through mass measurements determined by radial velocity. However, radial velocity measurements are not always possible for dimmer stars, due to photon noise (Santerne et al. 2013). For this reason, potential exoplanets around bright stars are able to be confirmed much more easily. The majority of *Kepler* observations are dim, and only about 5-10% are expected to be able to be verified through radial velocity measurements (Santerne et al. 2013).

Changes in the star's flux could also be caused by stellar variability such as spots or pulsations. There could also be instrumental issues that might appear to be periodic flux variations and might show up as false positives.

Another drawback of this method is that it is impossible to apply to systems that are not oriented edge on with respect to the Earth. If the system is perpendicular to the line of sight, any planets around the star will never pass in front of the star and never be seen. It is easiest to see planets with edge on orientation with little inclination between the planet's orbit and the star's orbit, since the planet will cover the star for longer and cover more of the star, resulting in deeper and longer reductions in flux and a more easily detectable transit. Nonetheless, the information able to be obtained from this method far outweighs the drawbacks, and it remains the most useful detection method at this time. With more missions being dedicated to detecting planets, there will be many more planets found via this technique.

1.1.4 *Kepler* Mission and *K2* Rebirth

Kepler launched in 2009 and was used to observe 150,000 stars near Cygnus and Lyra with a combination of continuous 30 minute and shorter 1 minute sam-

pling (Van Cleve and Caldwell 2009). *Kepler's* goal was to find exoplanets through transit detection, primarily terrestrial planets less than 2.5 Earth radii in the habitable zones around type F through M stars with apparent magnitudes between 9 and 15 (Van Cleve and Caldwell 2009). *Kepler* operated for a total of 4 years and found thousands of exoplanets, including numerous small planets, showing that planets smaller than the hot Jupiters typically found are plentiful.

After the first four years, two of *Kepler's* four stabilizing wheels broke, making targeting of specific stars more difficult. *Kepler* was reborn as *K2*, which operated from 2014 to 2018. *K2* observed stars along the ecliptic, since viewing this location subjects the spacecraft to less turbulence from solar wind pressure, allowing the remaining reaction wheels and thrusters to keep the craft at a relatively stable altitude (Van Cleve and Caldwell 2009). *K2* was initiated with a large number of goals in mind, mainly observing bright stars along the ecliptic. For exoplanet goals, *K2* aimed to determine how many large hot planets exist around bright stars in order to follow them up with transit spectroscopy to help improve what we know about characterization of exoplanet atmospheres, determine how many small planets are around bright, small stars in to better measure masses, densities, and composition of these planets, determine where habitable planets could exist around nearby M-dwarfs, and to determine if hot gas giants form around young stars close in, or if they migrate inward closer to their host stars through another mechanism¹.

Proposals for specific categories of stars could be requested. A multitude of projects have been approved, including a search for transiting planets around white dwarfs and for transiting planets around nearby stars were both proposed by Seth Redfield. A list of approved programs can be found on the *Kepler/K2*

¹<https://www.planetary.org/blogs/jason-davis/farewell-kepler.html>

website².

1.1.5 *TESS*

TESS (*Transiting Exoplanet Survey Satellite*) is the most recent NASA mission with the goal of detecting exoplanets. Unlike *Kepler*, *TESS* observes the whole sky. The mission is to observe 200,000 bright stars and identify whether or not they are orbited by planets definitive enough for follow-up with future missions. *TESS* observes the sky in 26 sectors: 13 in the northern hemisphere, 13 in southern hemisphere (Ricker et al. 2014). These sectors are shown in Figure 1.7.

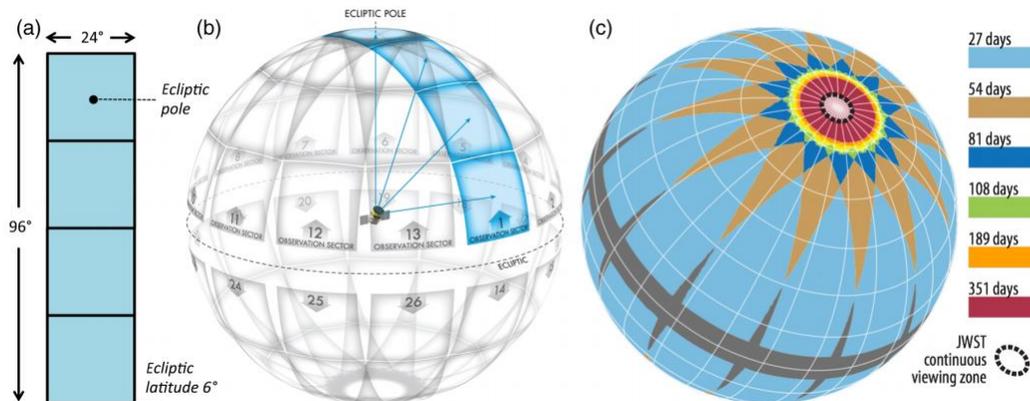


Figure 1.7: Diagram showing the observation sectors of TESS in order to thoroughly cover the whole sky. Image credit: Ricker et al. (2014)

This mission looks for brighter stars than *Kepler/K2*, since they are better targets for followup exoplanet studies. *TESS* saturates with stars for a center pixel brightness of $I_c = 7.5$ (Ricker et al. 2014), an improvement over *Kepler* and *K2* which saturated at a magnitude of $K_p = 11.3$ (Pope et al. 2019). However, there are still stars that saturate during observations by *TESS*, making traditional

²<https://keplerscience.arc.nasa.gov/k2-approved-programs.html>

data reduction methods ineffective. Figures 1.8 and 1.9 show the number of stars in both missions that are above the saturation limit.

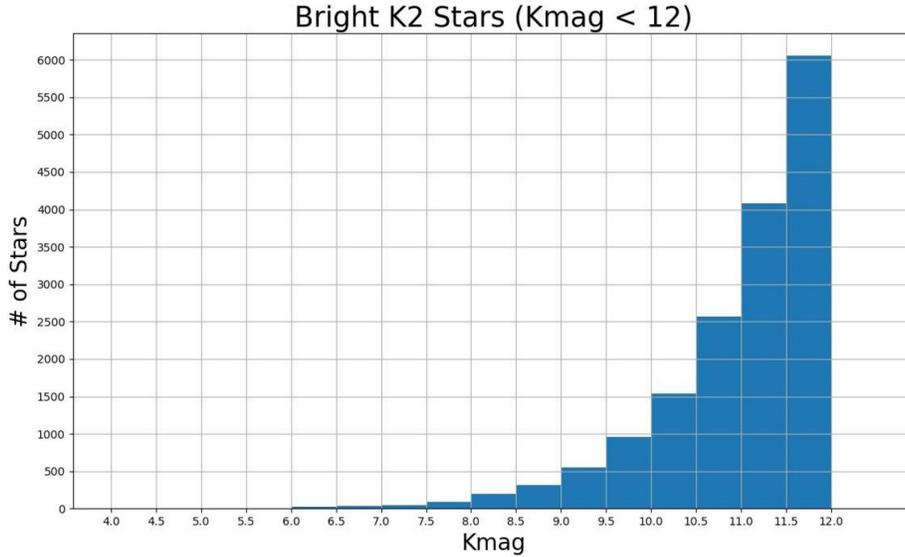


Figure 1.8: Histogram showing the number of bright stars per K2 magnitude above the saturation limit.

However, light curves can be easily recovered for stars a few magnitudes past the saturation limit. When a pixel saturates, the flux bleeds out of that pixel into neighboring pixels in the column. In extreme cases, this can cause bleed columns hundreds of pixels long. However, charge can be conserved for many bright stars, as long as the photometric aperture is large enough to contain all the flux (Ricker et al. 2014). If the CCD is large enough, even if a star causes saturation, as long as the flux does not bleed to the edge of the CCD, the charge is still conserved and the light curve can be recovered by adding up the pixels (Ricker et al. 2014). However, this is only possible with stars about four magnitudes past the single pixel saturation limit (Ricker et al. 2014). This would include stars up to $K_p = 7.3$ for *Kepler/K2* and $I_c = 3.5$ for *TESS*. However, there are still bright stars past this

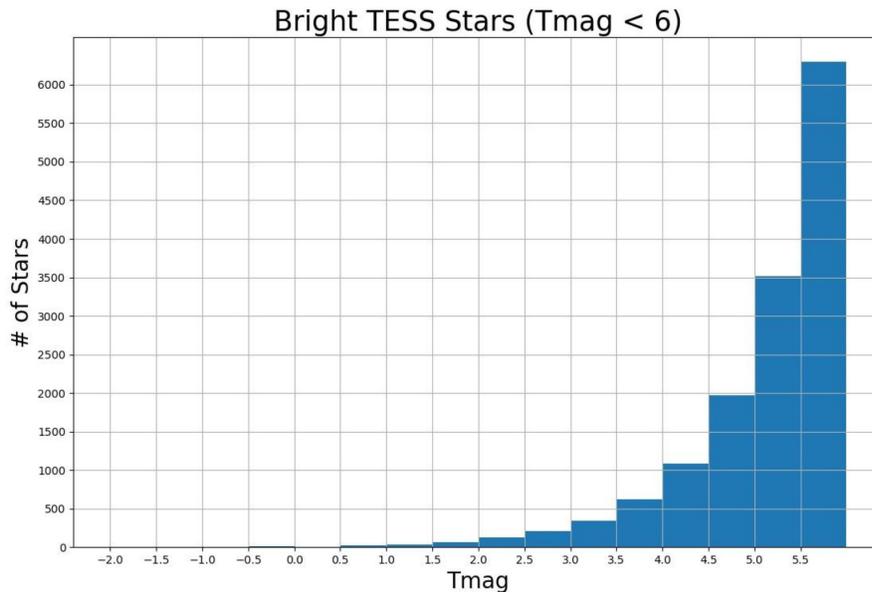


Figure 1.9: Histogram showing the number of bright stars per *TESS* magnitude above the saturation limit.

saturation threshold, especially for *K2* which observed a number of bright stars when its field of view shifted to the ecliptic, that do require special techniques. There are 125 stars observed by *Kepler* and 589 stars observed by *K2* that fall beyond this bright limit.

1.1.6 Future Missions

The biggest upcoming NASA mission is the *James Webb Space Telescope (JWST)*. This telescope is being created to improve what can be currently observed with *Hubble Space Telescope (HST)* by focusing on longer wavelengths in the infrared and increased sensitivity³. This telescope will enable follow-up, higher detail observations of promising exoplanet candidates from *K2* and *TESS*.

Another upcoming mission specifically searching for exoplanets is the *PLAn-*

³<https://jwst.nasa.gov/index.html>

etary Transits and Oscillations of stars (PLATO). This mission is part of the European Space Agency, and is to find and study a large amount of exoplanets, particularly terrestrial planets in the habitable zone around bright stars⁴. PLATO will have an on-board seismic activity monitor, enabling it to study host stars in more detail, determining characteristics such as age⁵. This mission will be able to perform follow-up studies on any promising exoplanets found around bright stars. PLATO is anticipated to launch in 2026.

1.2 Bright Stars

Bright stars are of interest as exoplanet hosts because they provide a high signal-to-noise that can be better used for follow up studies. When exoplanet transits are viewed, some of the spectrum of light that reaches the observer contains not only information about the star, but also about the planet's atmosphere. When the planets pass across a star, some of the light reaching the observer passes through the planet's atmosphere, while some of the light is blocked. This can result in spectroscopic features in the light curve indicating the presence of certain elements in the atmosphere of the planet. Brighter stars have higher flux and a higher signal-to-noise ratio that is ideal for identifying information about the planet's atmosphere amid the star's spectrum. For this reason, bright stars are excellent candidates for high-resolution spectroscopy and interferometry, which both require large amounts of photons (Pope et al. 2019). High-resolution spectroscopy is essential to follow-up studies of exoplanets like atmosphere characterization.

However, bright stars are problematic for traditional processing techniques because they saturate the CCD. Saturation occurs when there are so many incoming

⁴<https://sci.esa.int/web/plato>

⁵<https://sci.esa.int/web/plato>

photons from a bright source, the pixels that receive the energy and liberate electrons in response cannot hold all the charge. The pixels closest to the center of the star fill and electrons “bleed” over into neighboring columns (Pope et al. 2019). For bright stars, this “bleeding” will occur, filling more columns the brighter the star. As long as the CCD is large enough to contain the electrons, even if they “bleed” over into adjacent columns, the signal for the observation can still be retrieved through a process called simple aperture photometry (SAP). The SAP method requires adding up all the flux in a window around the bleed columns (Pope et al. 2019). Since charge is conserved, as long as all of the electrons are captured and do not bleed outside the bounds of the CCD, the signal can still be extracted.

Unfortunately, the SAP method only works for some bright stars. Specifically with *K2*, much larger apertures are required to perform SAP photometry on very bright stars, due to the increased spacecraft motion, which also made instrumental issues more difficult to correct (Pope et al. 2019). These systematic issues could be dealt with using the pipeline EVEREST 2.0, but even this pipeline could not work for the brightest stars that fully saturated the CCD, producing bleed columns all the way to the edge of the detector (Pope et al. 2019). Thus, another data processing technique had to be created to properly analyze these very bright stars.

1.2.1 Halophot

Halophot (short for “halo photometry”) is an algorithm that is able to extract light curves for bright stars even with high levels of saturation. It adds up flux in the pixels in the “halo” of scattered light around the saturated region of the star and is able to reconstruct photometry for the star through relative weighting

of pixels (Pope et al. 2019).

The total flux in the final light curve for an observation i (f_i) is constructed using the following equation:

$$f_i = \sum_{j=1}^M w_j p_{ij} \quad (1.4)$$

where w_j is the weight of a pixel with index j , p_{ij} is the flux in the pixel j at a specific observation i , and M is the number of total pixels. Therefore, the flux in each pixel in the final light curve for given observation, is equal to the sum of the flux in each pixel after weighting. If the weights were equal for each pixel, this would be identical to the SAP process previously described. Instead, the weighting is determined based on instrumental effects in *K2*. This is done by minimizing the normalized first-order total variation (TV) of flux, which minimizes the differences between successive observations during a time series (White et al. 2017). The equation is as follows:

$$TV = \frac{\sum_{i=1}^N |f_i - f_{i-1}|}{\sum_{i=1}^N f_i} \quad (1.5)$$

where N is the total number of observations. By minimizing TV and using this condition to weight pixels, the roll motion of *K2* can be minimized, without neglecting actual stellar variability (White et al. 2017).

Bright stars have large postage stamps, which end up capturing signal from numerous background sources. These background sources interfere with the weighting of the halo region. More stable, bright background stars can be more heavily weighted than pixels in the halo of the bright star which may be more variable. This can be seen in Figure 1.10. These background stars must be removed.

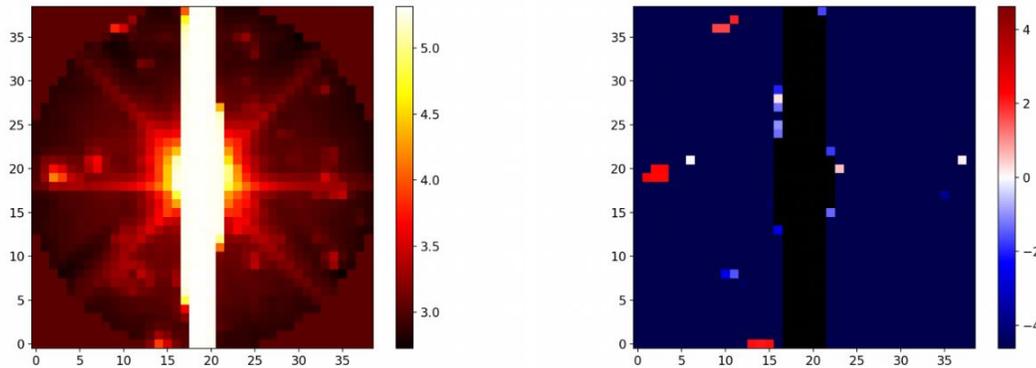


Figure 1.10: The left image shows a bright star with background stars. The right image shows the pixel weighting of different portions of the image. The background stars are more heavily weighted than most of the halo pixels. Image credit: Mireles (2019)

1.3 PSF Modeling

A point spread function (PSF) refers to the two dimensional light curve of a star when captured by a CCD, including any instrumental or observation effects that have altered the light.

For ground based telescopes such as *Sloan Digital Sky Survey* (SDSS), the PSF is typically dominated by atmospheric effects (Xin et al. 2018). However, for space based telescopes like *Kepler*, *K2*, and *TESS*, atmospheric effects are not a concern. The main factors influencing the PSF for these telescopes are instrumental properties.

Each telescope collects incoming flux differently and stores the incoming data in a different way, resulting in a different PSF. This PSF can also change with location and time. By stacking the observations of different stars from different locations, an averaged “shape” of the PSF can be created.

1.3.1 Double Gaussian

Previous attempts have been made to remove contaminating background stars from *Kepler* observations. PSFs are often modeled as double Gaussians around a point source, with the signal strongest in the center and tapering at the edges, like concentric circles of decreasing flux. The shape is composed of two Gaussian curves, where the luminosity is the highest in the center and tapers off like a bell curve. A Gaussian can be either axisymmetric, where the two Gaussians have the same width, or elongated where one axis is wider than the other. An example 2D Gaussian is shown in Figure 1.11.

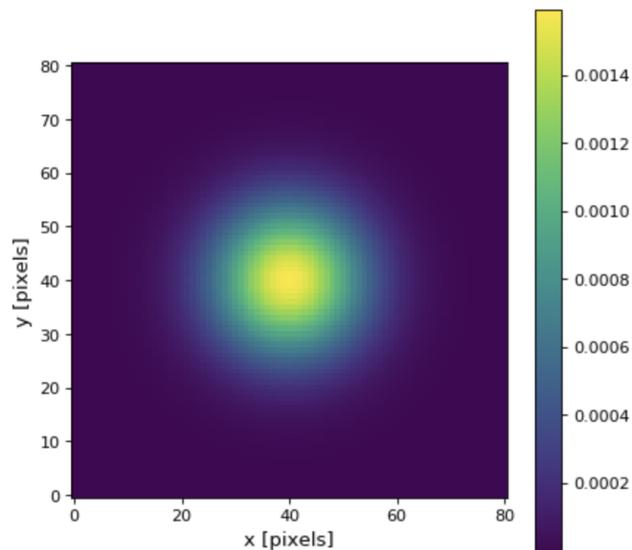


Figure 1.11: Diagram showing a 3D model of a Gaussian from <https://docs.astropy.org/en/stable/api/astropy.convolution.Gaussian2DKernel.html>

This model shows what would be expected for the PSF of a point source like a star. However, due to instrumental and observational effects, stars do not have a PSF perfectly aligned with a Gaussian shape. A 2D Gaussian works well as a

model for the center of a star, but not as well for the edges, especially for brighter stars (Mireles 2019). A fitting of a Gaussian for four different stellar observations with increasing brightness are shown in Figure 1.12 ($K_P = 16$), Figure 1.13 ($K_P = 14$), Figure 1.14 ($K_P = 12$), and Figure 1.15 ($K_P = 10$).

1.3.2 Halophot: Deathstar

Halophot has a function in the program that deletes background stars while processing bright saturated targets called `deathstar`. This function removes contaminated pixels from the background before the halo pixels are used to create the light curve. The pixels are chosen using the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm. The DBSCAN algorithm is able to efficiently identify clusters of any shape by starting at an arbitrary point and determining how many points are density-reachable (Ester et al. 1996). Any clusters other than the target star are selected by this method and separated using K2P2's watershed algorithm (Pope et al. 2019). The watershed algorithm, originally created to distinguish between different waterbeds for geology, can take the clusters of pixels and give them "relief values" based off their flux values, where negative is lower down and positive is higher up, allowing them to be separated similarly to the water basins (Lund et al. 2015).

Using these methods, the background stars can be identified, marked as separate regions of the image from the rest of the image, and removed from the "postage stamp". However this method has limited success subtracting contaminating stars near the brighter halo region of the star where it is more difficult to subtract background stars due to residual flux from the bright target star. The goal with PSF stacking is to be able to get a functional PSF model and subtract

a model of the stellar emission from the flux values where a background star is known to be. The contaminating stars can then be removed using their coordinates from *Gaia*, resulting in a more accurate light curve for the target star when using *halophot*.

1.3.3 PSF Model Creation

A new PSF model will be created by stacking multiple, single-star observations to get an average PSF for *K2*. It is important that these observations are uninfluenced by other stars, instrumental features, or saturated. By forming a PSF model from a conglomeration of observations, a more accurate model may be determined and used as a PSF characterization.

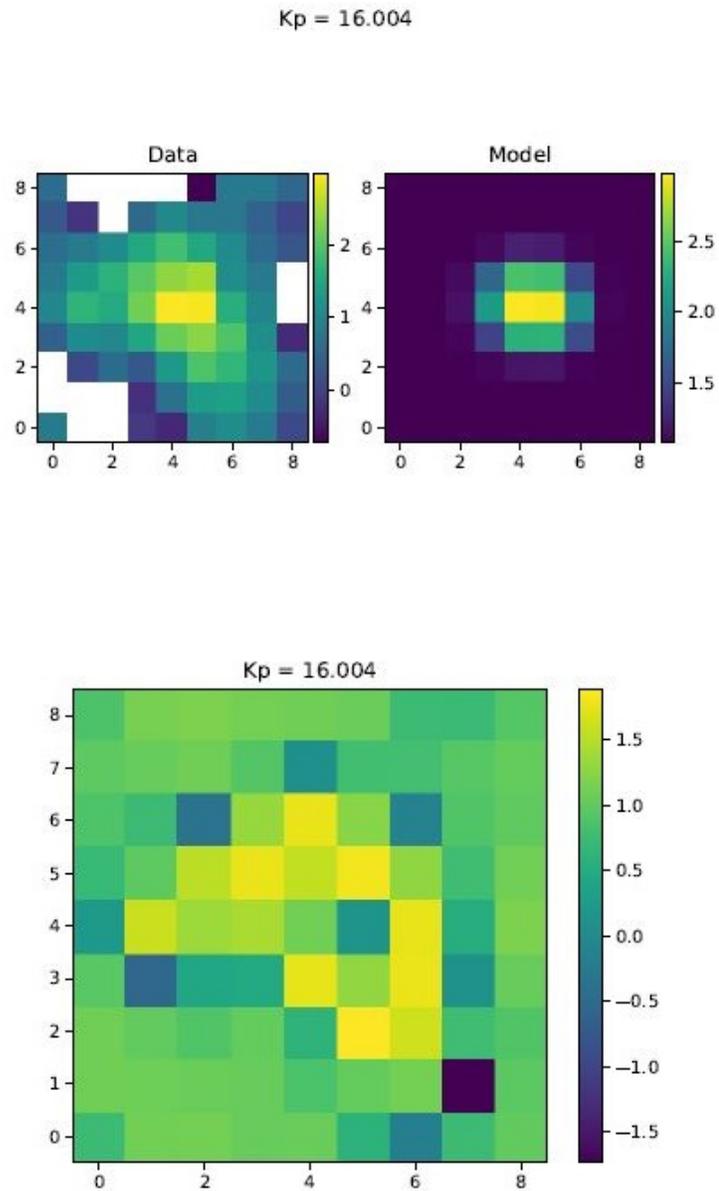


Figure 1.12: A figure showing a 2D Gaussian fitting to a $K2$ observation for a star of $K_P = 16$. The top two images show the observation (left) and the Gaussian model (right). The bottom image shows the fit through the residual. Image credit: Mireles (2019)

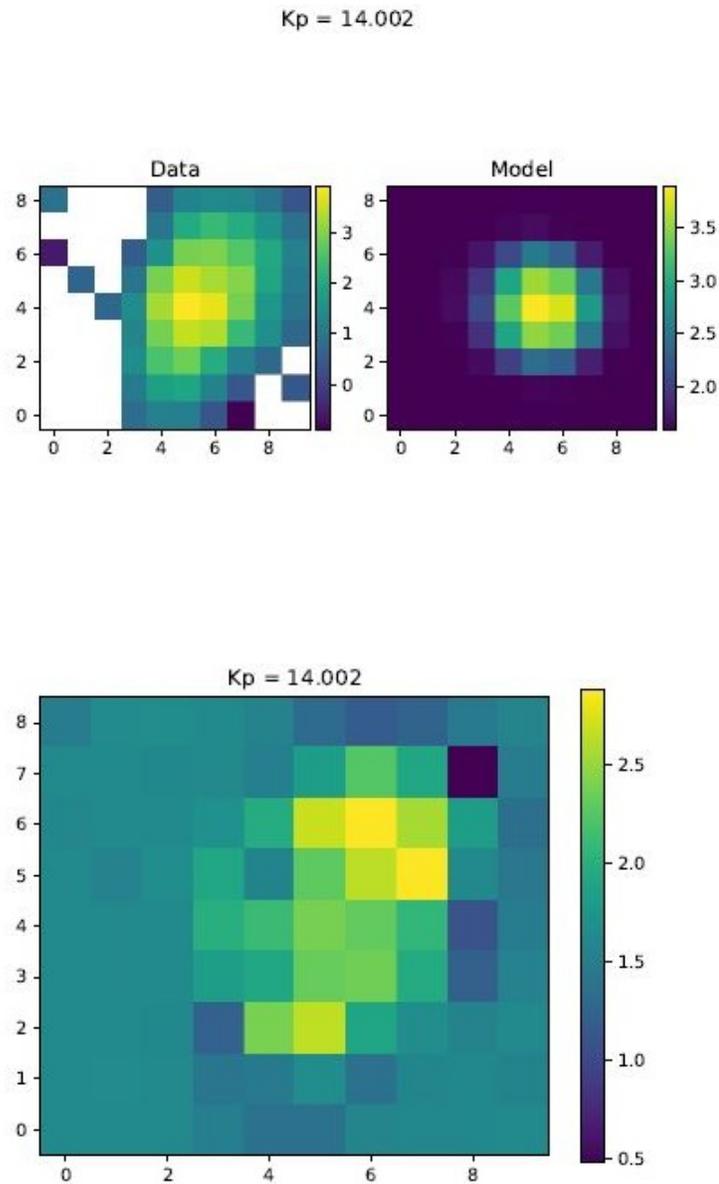


Figure 1.13: A figure showing a 2D Gaussian fitting to a K_2 observation for a star of $K_P = 14$. The top two images show the observation (left) and the Gaussian model (right). The bottom image shows the fit through the residual. Image credit: Mireles (2019)

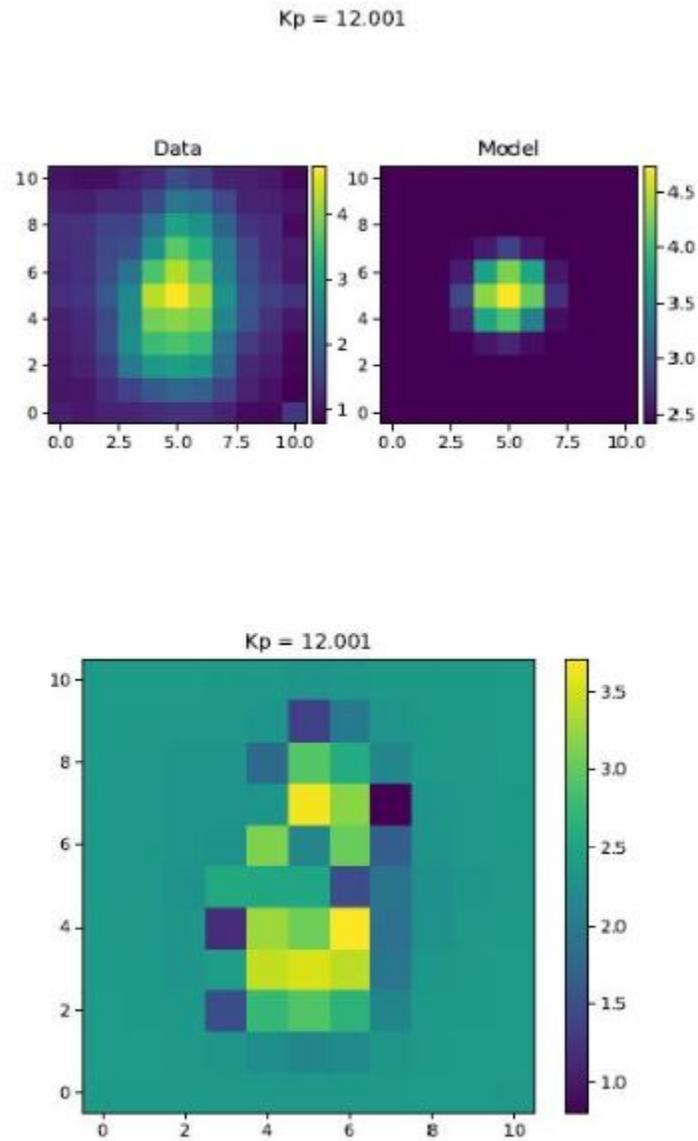


Figure 1.14: A figure showing a 2D Gaussian fitting to a $K2$ observation for a star of $K_p = 12$. The top two images show the observation (left) and the Gaussian model (right). The bottom image shows the fit through the residual. Image credit: Mireles (2019)

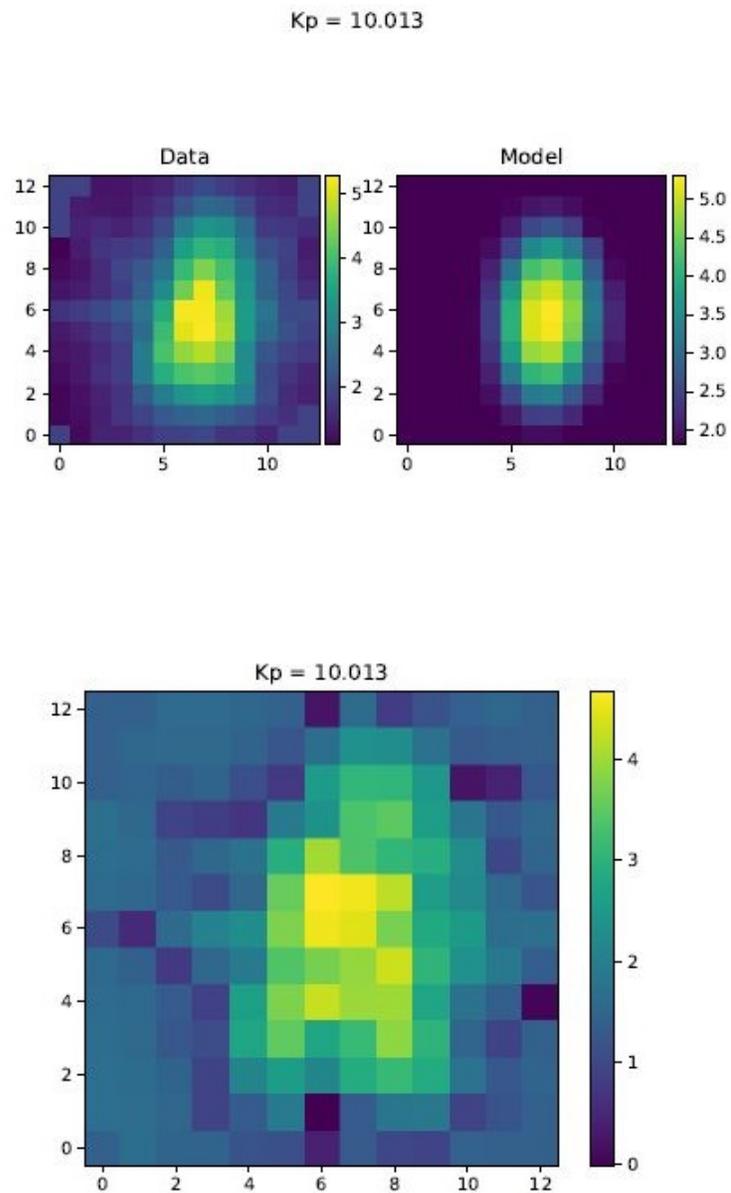


Figure 1.15: A figure showing a 2D Gaussian fitting to a $K2$ observation for a star of $K_P = 10$. The top two images show the observation (left) and the Gaussian model (right). The bottom image shows the fit through the residual. Image credit: Mireles (2019)

1.4 Goals for PSF Model

In this project, I will develop an average PSF model for six different channels in three different campaigns. This average PSF model will be created from all the uncontaminated, single star targets in each channel in each campaign. This will require identifying and removing any multi-star observations and any observations with other defects, instrumental or observational.

Once this PSF model is created, it will be compared to existing models. The two models for comparison are the previous PSF models from *Kepler*, called the *Kepler* PRF models, as well as a 2D Gaussian.

Based on the results of this comparison, the best fitting model can be used to try subtracting background stars from bright *K2* observations. This will hopefully result in a more accurate light curve for the bright stars, or for new light curves from the background stars being removed.

1.4.1 PSF Variance

In addition to the use of PSF model for subtracting background stars, the shape of the PSF and how it varies with time and location is also of interest. By making PSF stacks across different campaigns for the same instrument, the effects of time on PSF shape can be determined. By making PSF stacks across different channels for the same instrument, the effects of location on the field of view can also be determined. The focus is on PSFs from *K2* specifically, but the process can hopefully be applied to *TESS* and other future missions as well. Information can then be determined about differences in PSF characteristics between *K2* and *TESS* targets.

1.4.2 Exoplanets Around Bright Stars

Determining a more accurate PSF model of how *K2* modifies the shape of a star when capturing signal on the CCD can be used to better study bright stars. This new PSF model can be used to remove background stars from around bright stars in order to more accurately extract light curves using `halophot`.

With these modified light curves, exoplanet signals may arise that were not as apparent prior to finer subtraction. Exoplanet transits are sometimes small, especially around larger stars which are inherently brighter. It could be that there are exoplanets around these stars that are being missed due to residual signals from background stars affecting the light curve of the bright target star.

1.4.3 Exoplanets Around Background Stars

In each of the images containing a bright star, there are numerous background stars. When removing the background stars using the PSF, the light curve of each background star can also be evaluated. These background stars have not yet been evaluated for exoplanets since they are unintentional targets. If a more accurate PSF model can be created, not only will the background stars be removed more cleanly from the bright star observation, but their removal being more complete will result in a more complete light curve to study individually. These background stars will provide a numerous amount of new potential exoplanet hosts to evaluate by extracting and viewing their light curves.

Chapter 2

Data: *Kepler* vs. *K2*

2.1 The Instrument

Kepler, the predecessor to *K2*, was launched in May of 2009 with a planned operation of 3.5 years, which *Kepler* fulfilled (Van Cleve and Caldwell 2009). However, one reaction wheel failed in June 2012 and when a second reaction wheel failed in May 2013, the *Kepler* mission could no longer collect data in the same way (Van Cleve and Caldwell 2009). Without the two reaction wheels, *Kepler* could no longer accurately point towards the Cygnus-Lyra field. However, in November 2013 the *K2* mission was proposed to continue using the Kepler telescope to collect data, but on a modified path.

For the new observing region, the spacecraft would take advantage of light from the Sun and the fact that the solar panels were distributed evenly around the craft to use solar pressure to stabilize one axis (Cleve et al. 2015). This set up is shown in Figure 2.1.

By balancing the solar pressure, torque around the roll axis could be minimized. The remaining two reaction wheels were able to adjust the other two axes, allowing the telescope to point into the ecliptic plane for observations (Cleve et al. 2015). The pointing location of the telescope had to be moved approximately every 80 days due to changes in the Sun's position. The Sun could be too low in

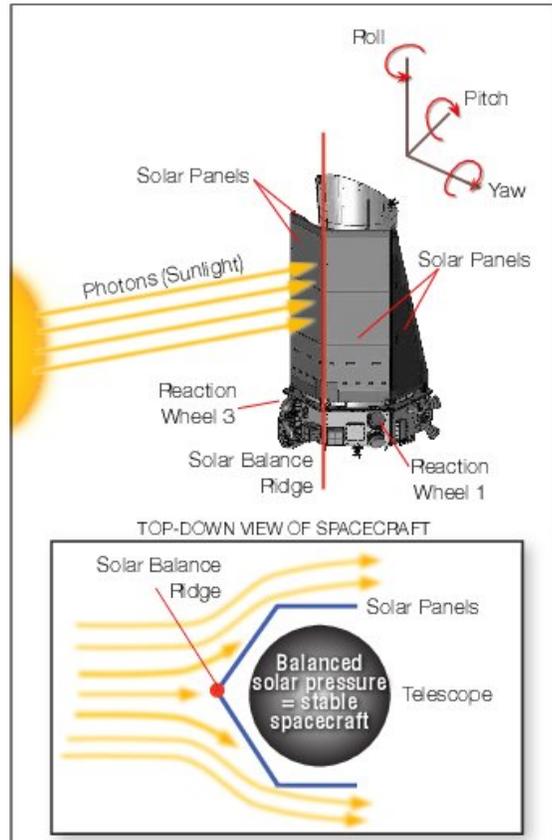


Figure 2.1: Diagram showing the orientation of *K2* to make use of solar pressure for stabilization. Image credit: Cleve et al. (2015)

the sky where not enough power was provided to the solar panels, or could be too high in the sky where it threatened the safety of the telescope (Cleve et al. 2015). For this reason, the field of view for *K2* shifts for each campaign. This process is shown in Figure 2.2. The *K2* mission became operational in May of 2014.

The only telescope on board *Kepler* was a 0.95 meter Schmidt telescope that transmitted the observed light onto a 94.6 million pixel CCD detector (Van Cleve and Caldwell 2009). The field of view is 16.1 degrees in diameter, with 115.6 square degrees of this area corresponding to active pixels (Van Cleve and Caldwell 2009). Active pixels are pixels capable of capturing flux, and with these dimensions they

occupied a bit more than half of the aperture, since the field of view had an area of about 200 square degrees. This information is identical for the *K2* mission since the instrument was the same, only the pointing changed.

However due to less stability in pointing, the precision for *K2* was not as high as the original *Kepler* mission. A comparison of the photometry precision between the two missions as a function of target magnitude is shown in Figure 2.3. Sensitivity to detection of 6-hr transits between the two missions was compared, though *Kepler's* targets were all quiet G-dwarfs and while *K2's* targets were of undetermined stellar type, though both sets of targets were randomly distributed across the field of view (Howell et al. 2014). The larger variation in *K2* targets is very likely the cause of the higher scatter in the *K2* data points.

The viewing time per field was also shorter than the original mission. Despite these drawbacks, *K2* still had much better sensitivity, field of-view, and continuous time coverage than ground-based telescopes (Howell et al. 2014). Even with less stable pointing and precision than *Kepler*, *K2* was still a space-based telescope and able to get observations without having to deal with the atmospheric effects that ground-based telescopes would.

This regeneration of the original mission allowed for important data for discovery of exoplanets to be captured.

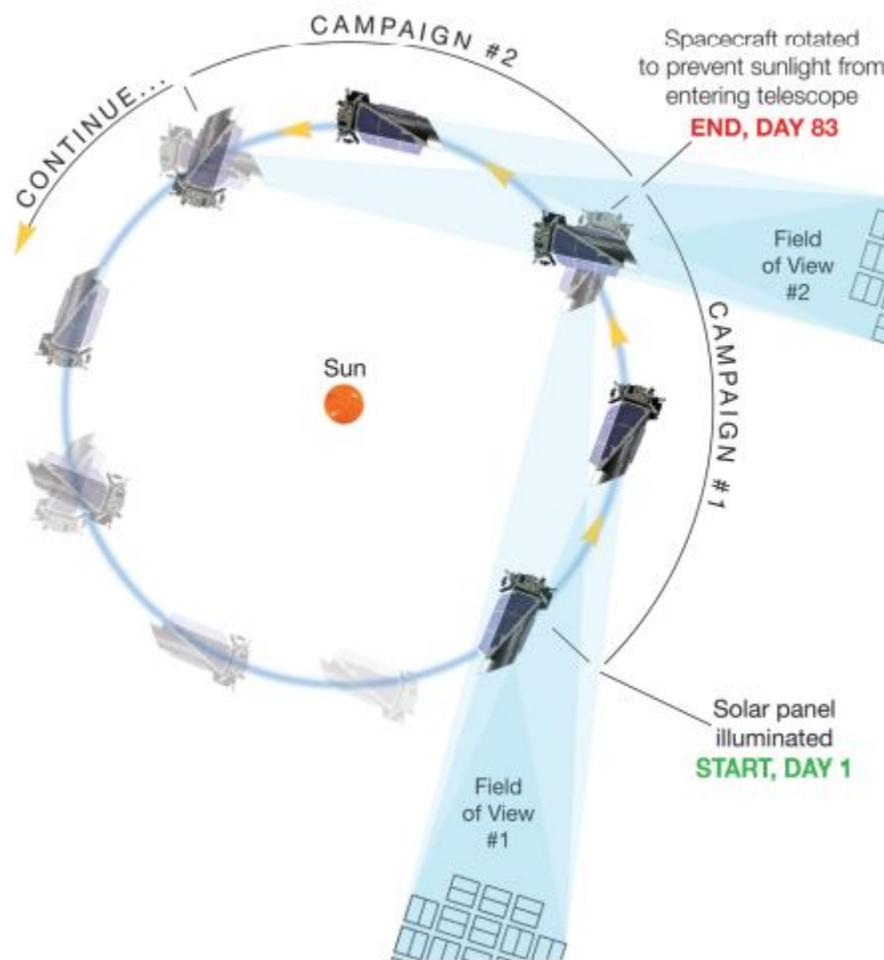


Figure 2.2: Diagram showing how the telescope is rotated for each new campaign in K2, in order to avoid direct sunlight entering the CCD. Image credit: Howell et al. (2014)

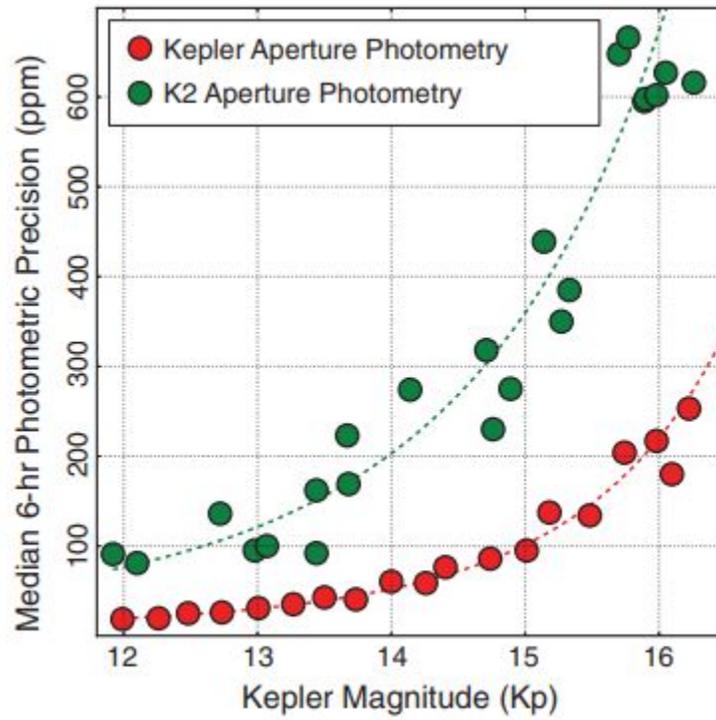


Figure 2.3: *K2* test of precision showing that *K2*'s precision is within 3-4 times the precision of *Kepler*. Image credit: Howell et al. (2014)

2.2 Observations and Data Processing

Kepler downloaded data once a month at a rate of only 550 kilobytes per second (Libralato et al. 2015). Due to this slow speed, *Kepler* performed preliminary processing on most of the data before sending it to Earth in order to reduce the amount of information being sent. This reduction process included selecting a small “stamp”, usually a square of only a few pixels, around each target every 6.02 seconds (Libralato et al. 2015). Bright sources required larger postage stamps, due to the larger flux. Sometimes, especially with saturated examples, these bright stars can lead to postage stamps hundreds of pixels long. An example is shown in Figure 2.4.

These postage stamps were added together for a total integration time of 29 minutes every 270 exposures (Libralato et al. 2015). These almost half hour integration time observations are referred to as “long-cadence stamps”. Shorter integration times of 1-minute from 9 exposures were also created for some targets and are called “short-cadence stamps” (Libralato et al. 2015). A collection of these short-cadence and long-cadence stamps is called a target-pixel file (TPF). These TPFs can be downloaded from the MAST (Mikulski Archive for Space Telescopes) website¹ by a specified campaign and further constrained by a variety of other markers such as brightness, location, channel, etc. by sorting a table of TPFs for a specific campaign and individually selecting the files for download.

Each pixel corresponds to 3.98 arcsecs for the *Kepler* telescope (Jenkins 2017). *Kepler* has 1024 preloaded aperture masks that are selected based on the smallest mask that can fit all the necessary pixels for an observation (Jenkins 2017). Some of the brighter targets may be fit with more than one mask (Jenkins 2017). This

¹https://archive.stsci.edu/k2/search_retrieve.html

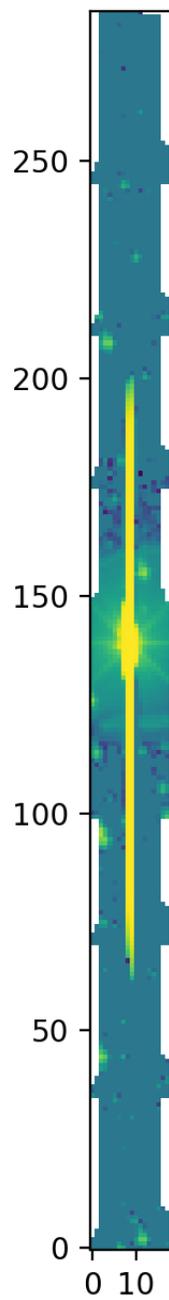


Figure 2.4: Example bright star, with saturation causing “bleed” columns hundreds of pixels long.

is because brighter targets need more pixels to contain all the flux. These masks led to the selection of the “postage stamps” described above.

On average a PSF for a star is several pixels across, except at the center of the field of view where a PSF might be up to 7 or so pixels across (Libralato et al. 2015). The result is relatively low resolution images. The low resolution is helpful for *K2* since the pixels over the course of time series for *K2* change more than they did for *Kepler* due to reduced pointing precision (Cleve et al. 2015). Due to consecutive pixels differing more dramatically, more space is required to encode each pixel. Lower resolution is helpful to reduce the space required, and also to minimize the changes in signal.

Despite the low resolution, features of the PSF are still undersampled throughout the field of view (Libralato et al. 2015). Undersampling refers to there not being enough pixels to properly capture the full resolution of the PSF.

Data for the *K2* mission was very similar to data from the *Kepler* mission, apart from the reduction in precision. The cadence time for *K2* was identical to the times stated above for *Kepler* for both long and short cadences (Van Cleve and Mighell 2019). Data was downloaded and processed in the same way. Observations were marked by “campaign”, a cycle of about 81 days with differing fields of view (Van Cleve and Mighell 2019). These changes in field of view are shown in Figure 2.8. *K2* had a total of twenty campaigns, starting at campaign 0 and ending at campaign 19. The data for campaigns 9, 10, and 11 were each processed in two separate portions, referred to as part a and part b for each campaign (Van Cleve and Mighell 2019). These campaigns had to be processed in two parts for various reasons. Campaign 9 had to be downloaded in two chunks due to the high concentration of stars, as data was only downloaded at the end of

a campaign². Campaigns 10 and 11 had to be separated into two parts due to pointing errors at the start of each campaign that needed to be corrected partway through³. Other than those campaigns, the campaigns were all processed as single segments.

One major difference between the *Kepler* and *K2* data sets was the amount of stars per square degree in the observing fields. In *Kepler*, the stellar density per square degree was relatively constant due to focusing on the same field in the sky. However, for *K2* the stellar density varied significantly (Van Cleve and Mighell 2019). Unlike *Kepler*, *K2* was not pointed at a continuous location. Instead, *K2* was set to observe along the orbital path of the telescope, which is near the ecliptic. Along this path, the field of view is constantly changing, resulting in large variations of stellar density. Additionally, within the ecliptic, there are portions of very high stellar density. The variance in stellar density can be seen in Figure 2.5.

This variance in stellar density led to a wider variety of stellar observations, but also more crowded fields of view. When looking at *K2* observations, some campaigns have very dense stellar fields, particularly campaign 7, 9, 10, and 11⁴. This high stellar density is detrimental to finding a collective PSF for a single star, as background contamination from multiple stars in the field of view for an observation can lead to stray light altering the PSF. These portions of the mission would most benefit from having a functional PSF model to remove field stars.

²<https://keplerscience.arc.nasa.gov/k2-data-release-notes.html#k2-campaign-9>

³<https://keplerscience.arc.nasa.gov/k2-data-release-notes.html>

⁴<https://keplerscience.arc.nasa.gov/k2-campaign-7-data-available.html>

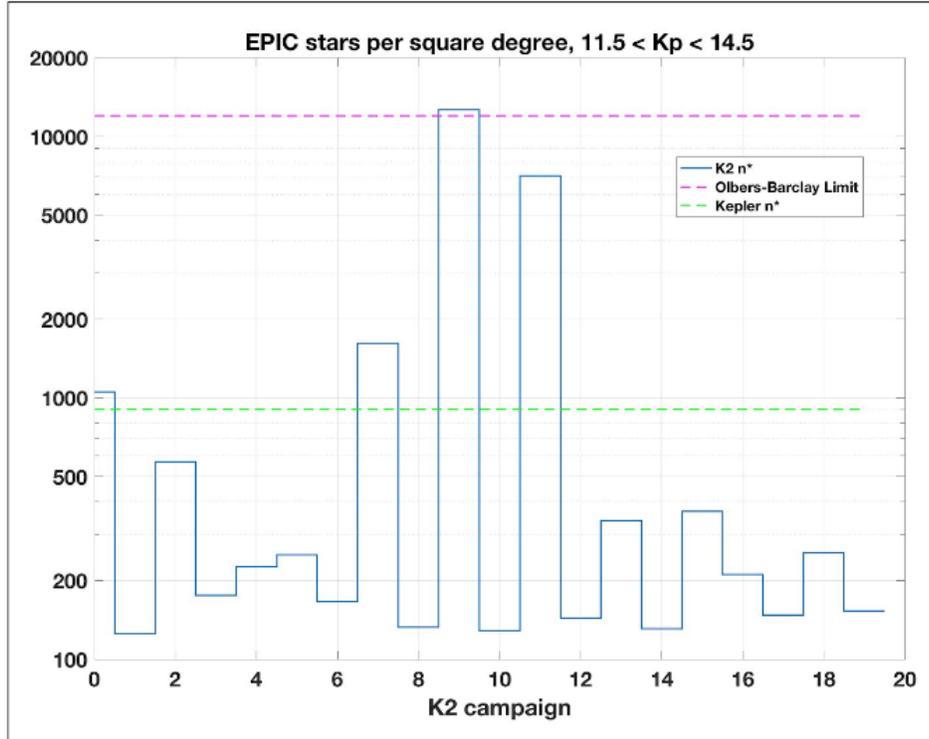


Figure 2.5: The variation in the stellar density per square degree in *K2*'s field of view. Due to the changing field of view along the ecliptic there is a wide range of stellar densities. Image credit: Van Cleve and Mighell (2019)

2.3 Field of View

Kepler's field of view includes 21 modules, each made of two 1024×2048 CCDs either side by side vertically or stacked horizontally, for a total of 42 CCDs (Libralato et al. 2015). The configuration of the CCDs, plus the location of the field of view for *Kepler* is shown in Figure 2.6. Each of the CCDs is read out as two separate 1024×1024 channels, for a total of 84 channels (Libralato et al. 2015). During the *Kepler* mission, one module (MOD-3) failed in January 2010, followed by a second module (MOD-7) in February 2014 (Libralato et al. 2015). At the end of the *Kepler* mission, only 38 CCDs with 76 channels were still operational.

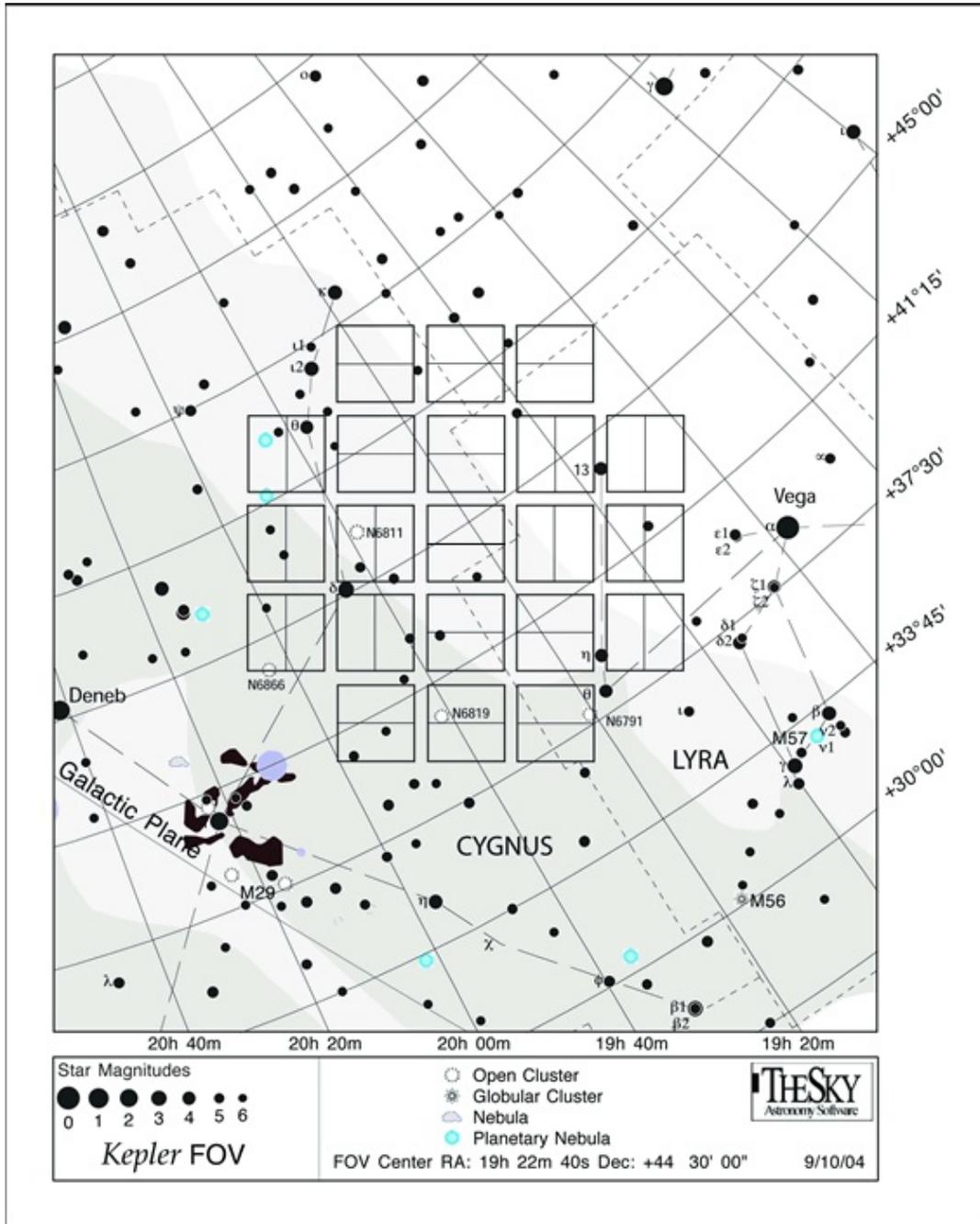


Figure 2.6: *Kepler*'s field of view at the start of the missions. *Kepler* was focused on one location in the northern circumpolar region between the constellations of Lyra and Cygnus. Image credit: Van Cleve and Caldwell (2009)

During 2016, in the middle of campaign 10 of *K2*, another module (MOD-4) failed (Van Cleve and Mighell 2019). This left *K2* with 36 CCDs with 72 channels, which remained operational throughout the remainder of the *K2* campaigns. All of the modules lost are greyed out in Figure 2.7. This figure also shows the readout number, channel number and module for each portion of *K2*'s focal plane.

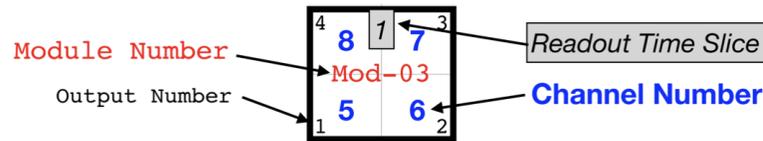
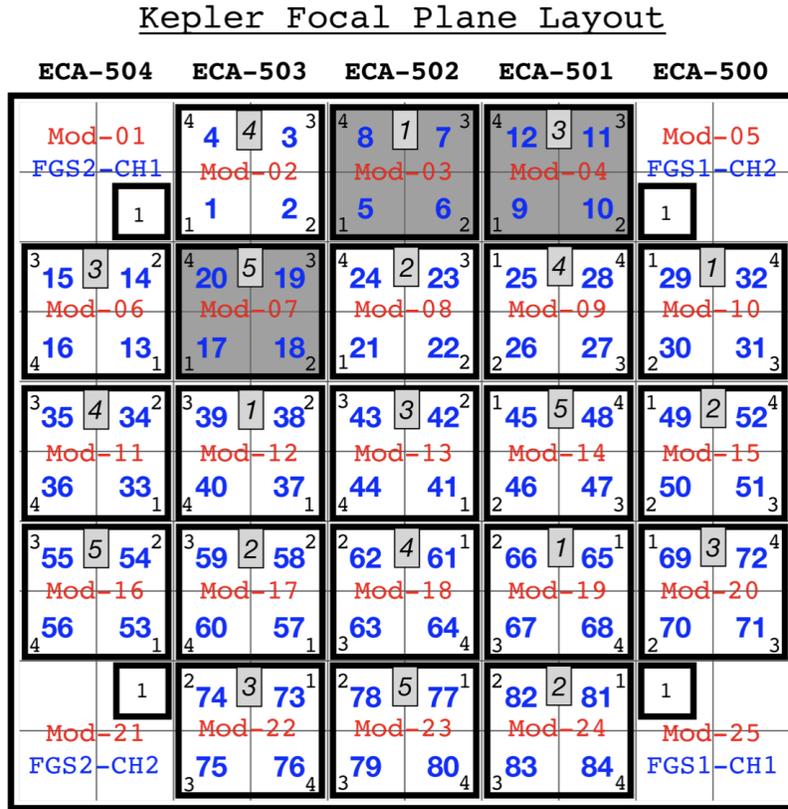


Figure 2.7: *K2*'s focal plane showing each module, channel, and readout number. The greyed out modules were not operational for the full duration of the mission. Image credit: Johnson (2015)

Kepler was chosen to point between the constellations of Cygnus and Lyra in the northern circumpolar region at a ecliptic latitude of 65 degrees (Van Cleve and Caldwell 2009). This latitude was chosen because of its proximity to the plane of the Milky Way so that stars would be plentiful, but also because it was pointed enough above the galactic plane that there would not be too many bright giants visible from large distances in the field of view (Van Cleve and Caldwell 2009). This higher latitude also permitted *Kepler* to view continuously with the attached sunshade (Van Cleve and Caldwell 2009).

For the duration of the *Kepler* mission, the spacecraft remained fixed observing at this location. However, every 93 days (referred to as a quarter), *Kepler* rotates on its optical axis by 90 degrees in order to keep the Sun on the side of the solar panels and the cooling radiator pointed towards deep space (Van Cleve and Caldwell 2009). This process is referred to as a “quarterly roll”. Due to this process, even though the same area in the sky was observed over the full course of the *Kepler* mission, a specific star would be viewed in four different locations on the focal plane, depending on the season since *Kepler* orbited on a heliocentric orbit (Van Cleve and Caldwell 2009).

However, once the two reaction wheels failed and the mission was reborn as *K2*, the field of view changed regularly. With only two wheels left to control motion, the telescope could only be controlled on two axes, allowing the field of view to wander along the other axis. *K2* was constrained to the ecliptic plane, rather than to a specific viewing location. This resulted in the varying field of view shown below in Figure 2.8.

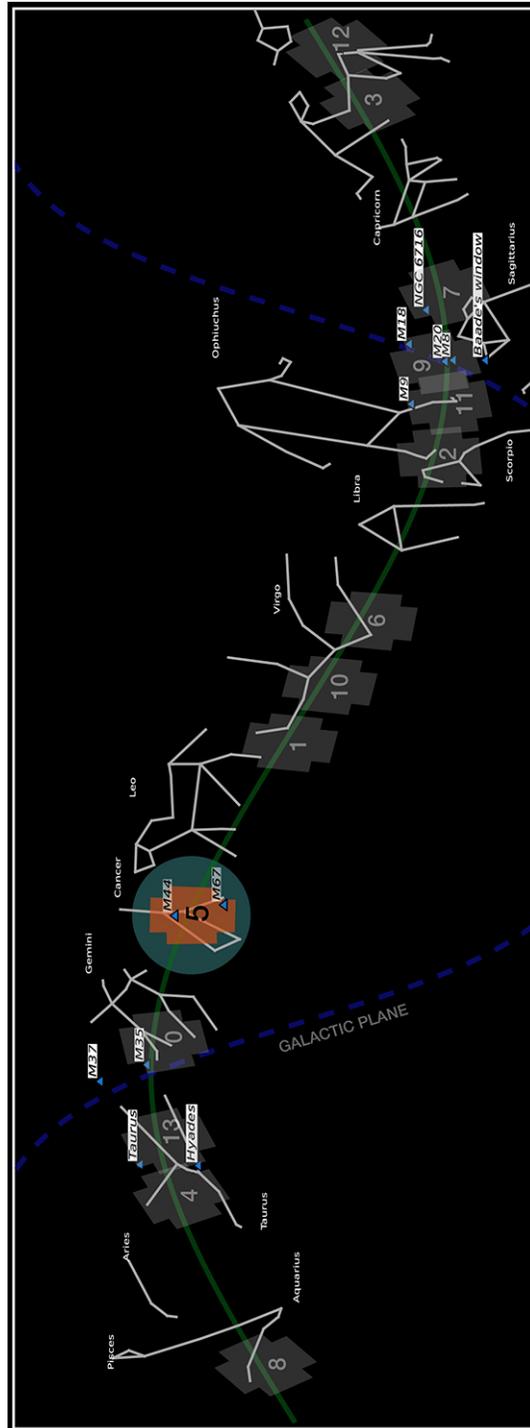


Figure 2.8: *K2*'s field of view (Johnson 2015)

2.4 Motion

For *Kepler*, the only motion that had to be corrected for was the reorientation of the spacecraft in regards to the Sun, referred to as quarterly rolling. This process took place regularly throughout the mission at intervals of 90 days. However, since the field of view remained fixed on the same location, despite the rotational orientation shifting four times a year, the target and guide stars for each quarter of the mission could be transferred from one CCD to the next (Van Cleve and Caldwell 2009). Targets were selected prior to launch from the *Kepler* Input Catalogue (KIC) which remained stable over the course of the missions (Van Cleve and Caldwell 2009).

However, when two of the reaction wheels broke and *Kepler* was reborn as *K2* the ability to focus on a single area of the sky was lost. Rather than locking on one location, *K2* slowly rotated in a circle near the ecliptic, rotating about 81 degrees every 80 days, initiating the beginning of a new campaign (Van Cleve and Caldwell 2009). Due to this constant change in FOV, target management had to be modified for each campaign in *K2*, resulting in more unique criteria for each portion of the mission. For *K2*, targets were selected from the Ecliptic Plane Input Catalogue (EPIC) which grew with each additional campaign (Van Cleve and Caldwell 2009). With each campaign a FOV had to be defined, targets within that field of view had to be selected, and aperture pixels that fit the selected targets had to be chosen (Van Cleve and Caldwell 2009). Due to this, no two campaigns in *K2* are alike, and the pipeline had to be rewritten to accommodate this continuous change.

Additionally, because of this constant adjustment in FOV, each module in *K2* has only one guide star, rather than the 10 guide stars in each module present

during the *Kepler* mission (Van Cleve and Mighell 2019). Reliable guide stars were so sparse in some fields that the requirement of including one actually influenced the FOV for some campaigns (Van Cleve and Mighell 2019).

2.4.1 *K2* Thruster Firing

In addition to the changing field of view, *K2* also experienced thruster firing events. Solar radiation pressure was able to keep the rotational axis of the craft in an “unstable equilibrium” with only slight rotational errors (or roll errors) along the boresight. The boresight refers to the optical axis, where the telescope is directing observations throughout its orbit. However, without control of the third reaction wheel, the thrusters on had to be fired periodically to correct these slight errors (Van Cleve and Mighell 2019). For *Kepler*, the boresight is pointed approximately into the ecliptic, thus resulting in observations along the ecliptic. The thruster firings are very brief (20–100 milliseconds) and occur after every twelfth long-cadence observation (which is approximately every 6 hours) (Van Cleve and Mighell 2019). These firing events, though scheduled between long-cadence observations, can cause quality flags on *K2* data. Quality flags indicate any issues that may effect data. The thruster firing events cause a characteristic sawtooth pattern shown in Figure 2.9. These firing events are corrected as part of the *K2* data reduction pipelines.

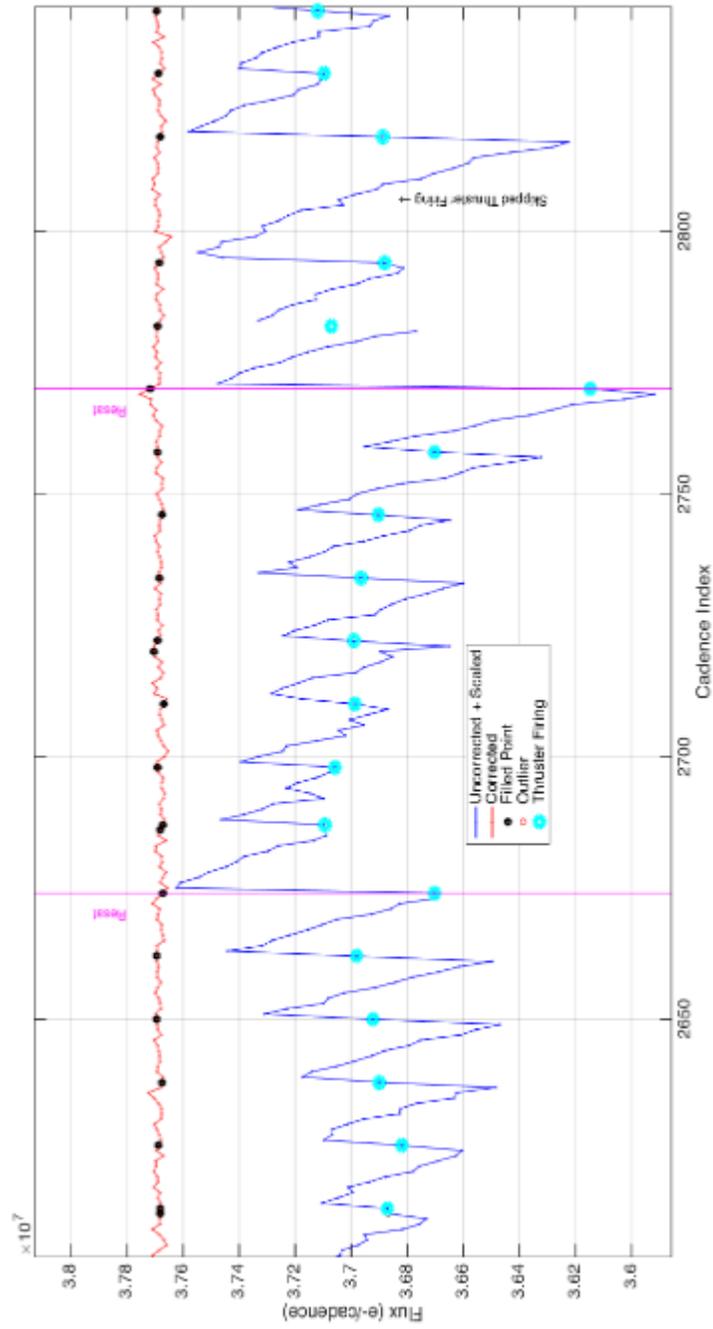


Figure 2.9: Image showing the characteristic sawtooth pattern in *K2* observations due to thruster firing events. Image credit: (Cleve et al. 2015)

2.5 Summary of *K2* vs. *Kepler* and the Effect on PSF

Kepler and *K2* both use the same telescope for observations and as a result, much of the instrumental features are the same. However, due to the loss of a reaction wheel, *K2* can only be controlled on two axes while the third axis, the “rolling” axis can only be approximately stabilized with solar pressure and adjusting thruster firing events. The reduction in pointing control leads to a decrease in precision and resolution since pixels change more throughout a flux time series, requiring more hard-drive space for each pixel to be stored.

Another issue is the ever-changing field of view for each *K2* campaign. *Kepler* pointed at a single area of the sky, but *K2* has a new field of view about every 80 days. Due to these changes, *K2* may require a PSF model for each campaign, a feature that *Kepler* does not have. The *Kepler*’s PSF was relatively well-defined, due to the consistent field of view. However, a PSF for *K2* is harder to accurately model, due to the effects described above. It is likely the PSF for *K2* will differ not only from *Kepler*’s PSF, but may differ within each *K2* campaign.

Chapter 3

Methodology

3.1 PRF vs. PSF

There are two different but similar acronyms: PRF (pixel response function) and PSF (point spread function). A PSF refers to how a CCD interprets the shape of a point source. Ideally, a point source would be characterized as a double Gaussian, with the signal slowly and symmetrically dropping off with increasing distance from the center. However, this is not the exact shape that is seen due to systematic effects. A PRF refers to how a CCD interprets the shape of a point source, coupled with the “spacecraft pointing jitter” and other mechanical effects (Bryson et al. 2010). This PRF has been modeled for *Kepler* and predicts what flux a star will have on any portion of the field of view, depending on the channel and location of the star. The PRF is different depending on the channel of the observation with channels towards the center of the field of view having a more Gaussian shape while channels towards the edges of the field of view have a more distorted shape. This effect is shown in Figure 3.1.

There are two methods of generation for the *Kepler* PRF. The `KeplerPRF` model can be used, which allows the user to specify the flux, center position, width, and rotation angle of the observed target (Bryson et al. 2010). In comparison, the `SimpleKeplerPRF` model only requires the flux and center position for the ob-

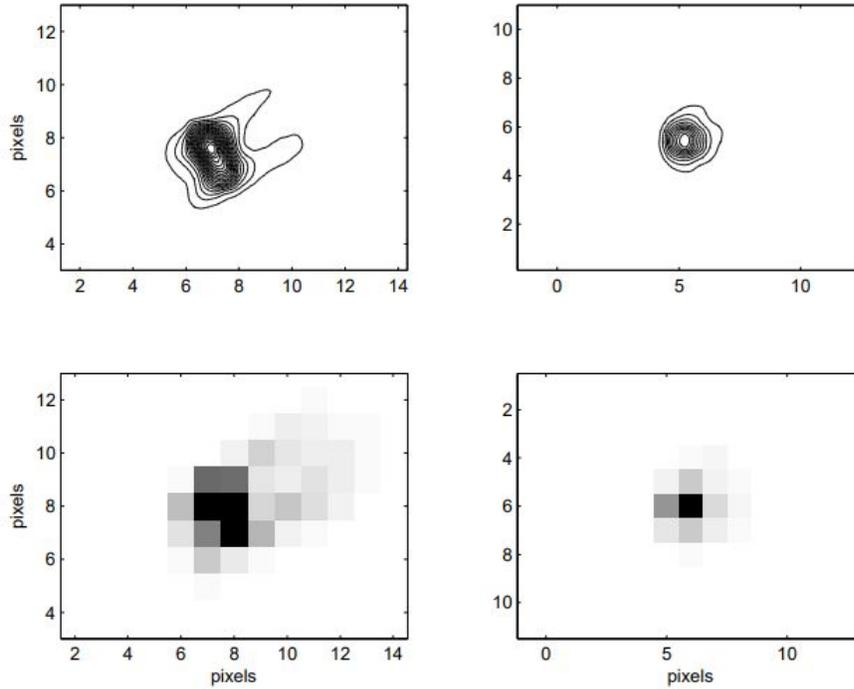


Figure 3.1: Diagram showing models of the PRF (on the top) and the PRF degraded to *Kepler* resolution (on the bottom). The right images are a location near the center of the CCD and the left images are a location near the edge of the CCD. Image credit: Bryson et al. (2010)

servation, with values fixed at 1.0 for width and 0 for rotation angle (Bryson et al. 2010). Since the PSF stacks being created are averages of numerous observations, the model with fixed rotation will be used for a more basic model.

PRF models only exist for *Kepler*. There are not separate PRF models for *K2*. However, since the same instrument is involved, only a difference in field of view and viewing conditions, the PRF models from *Kepler* should be similar enough to be applied to *K2* observations. These PRF models from *Kepler* will be compared with the new *K2* PSF model created from stacking observations and compared across each channel.

3.2 Contaminating Stars

When modeling the PSF, it is important to only use observations that contain single stars. Excess contaminating stars can lead to an inaccurate shape, since the flux from more than one star is contributing.

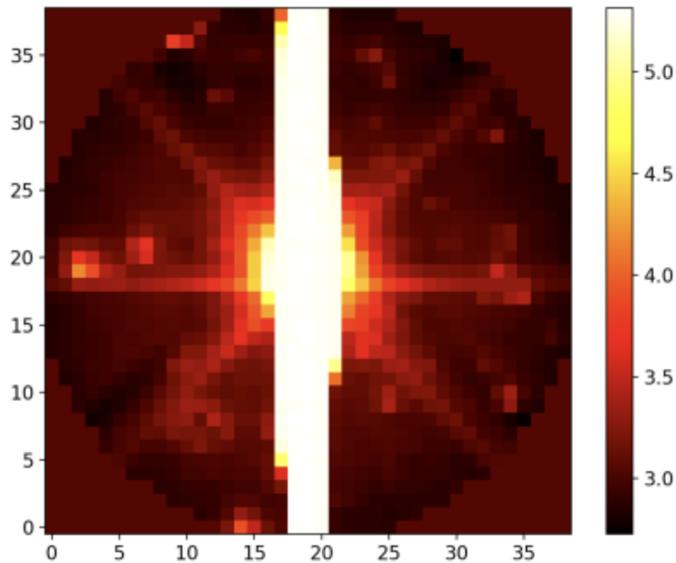


Figure 3.2: Diagram showing a flux slice for a bright star with contaminating stars in the background (Mireles 2019)

In Figure 3.2, a saturated bright star is shown, with several contaminating stars in the background of the image. These background stars must be removed before processing the bright star with `halophot`, in order to extract flux not contributed by the bright star.

In creating this average PSF for subtraction purposes, it is important to not use any stars with significant contamination from other sources as this will alter the shape of the PSF. Thus, any observations with multiple stars or other sources of contamination must be removed prior to using the remaining observations to

create an average PSF.

This was accomplished in two steps. The first screening involved using the Python package `astroquery` to search the *Gaia* database for all stars within 120 arcseconds of the given right ascension and declination values for each observation. Any observations that returned more than one star through *Gaia* were removed from the stacking sample. In Figure 3.3, an observation with two stars in it is shown:

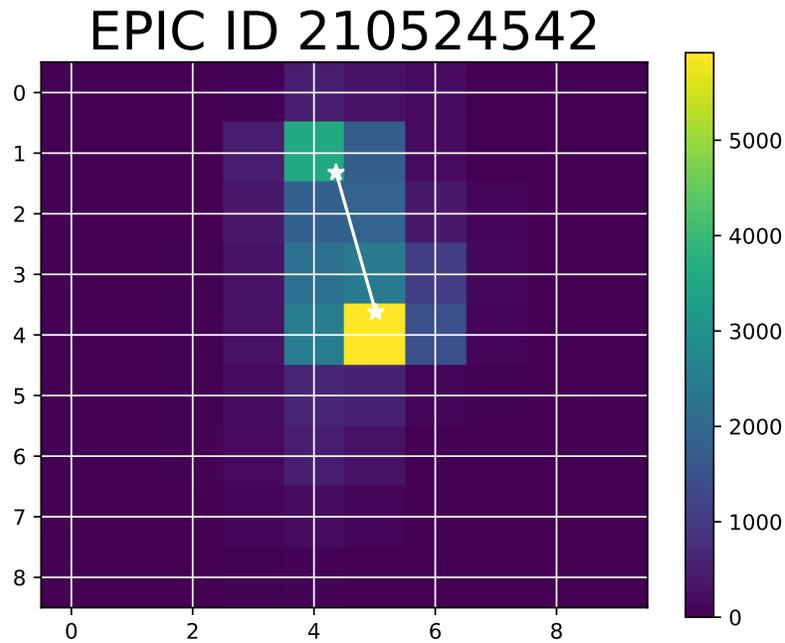


Figure 3.3: A flux snapshot showing an observation of two stars. The location of each star, according to the *Gaia* database, is marked with a white star.

These multi-star images needed to be removed in order to ensure that the average PSF being created was the result of stacking numerous single-star observations to get the average PSF of a single star, not the PSF of multiple stars blended into one image.

However, this method was only able to catch some sources of contamination. Some of the remaining images still contained background stars not known to *Gaia*, as well as other sources of contamination like saturation from outside of the postage stamp. In order to remove any remaining contaminated observations, a flux still from each image was visually inspected to look for any contamination missed by *Gaia* in the process above, or any other instrumental affects that might cause undue variation to the PSF shape. This was done by visually inspecting a flux frame for each observation and individually removing each additional contaminated observation from the data pool.

The remaining contaminated observations fell into one of five categories. The first type were multi-star observations not caught by *Gaia* (Figure 3.4). There were also stars that appeared to have some sort of contamination in half the image, possibly due to a bright star off the edge of the postage stamp (Figure 3.5). There were some bright, saturated stars that exhibited column bleeding (Figure 3.6). Some observations had more widespread, uniform saturation (Figure 3.7). Finally, there were stars at the edges of the postage stamp with a large portion of their flux cut off, rendering their PSF contribution unusable (Figure 3.8).

The remaining observations were stacked by channel. This was done by inserting each observation into an empty two-dimensional array, aligning each center with the center of the new array. The center location was determined using the right ascension and declination values for each file from the WCS (world coordinate system) coordinates and performing a *Gaia* search for stars in a specified radius around these values. Once a likely target was found, the coordinates from *Gaia* were converted back into pixel values. The *Gaia* coordinates for each star's center were more accurate than the coordinates given in the header. The comparison of the *Gaia* coordinates to the header coordinates is shown in Figure 3.9.

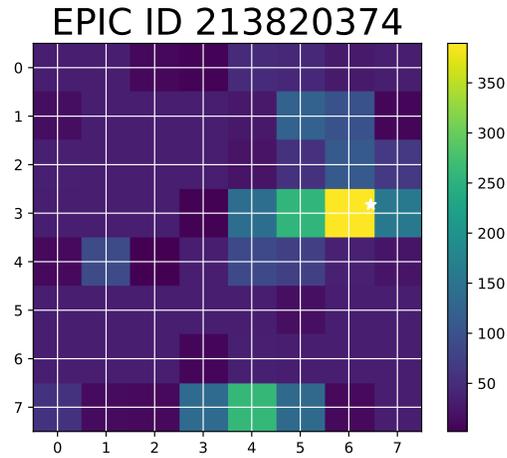


Figure 3.4: The figure above shows a presumably multi-star observation that did not have any background stars located by *Gaia*.

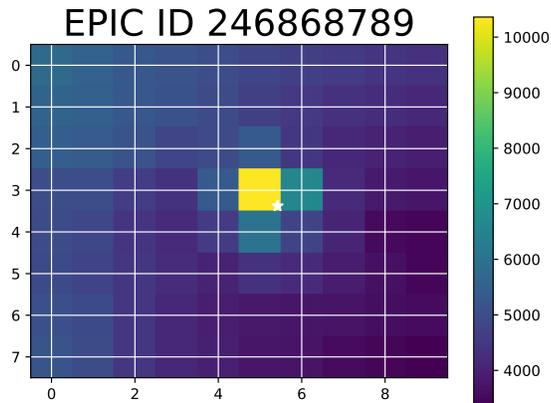


Figure 3.5: The figure above shows a single star observation with some sort of external light contamination towards the upper left of the postage stamp, most likely a bright star off screen.

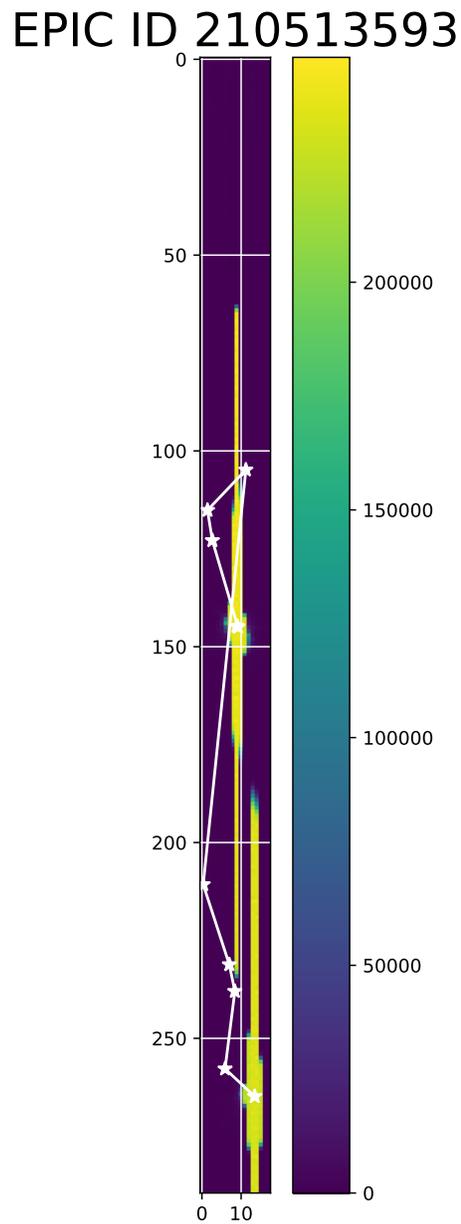


Figure 3.6: The figure above shows an observation where there are two bright stars, both exhibiting column bleeding.

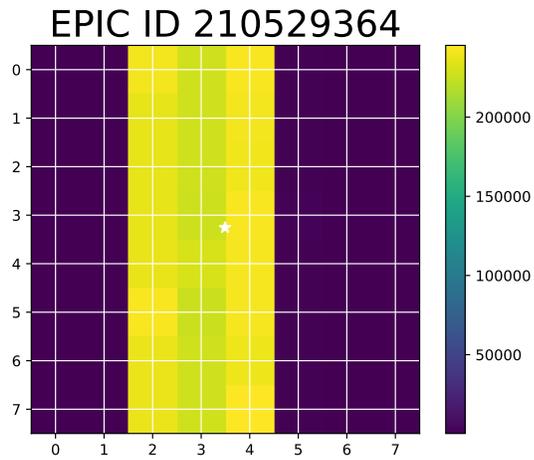


Figure 3.7: The figure above shows an observation where the target of interest may be a bright star, with saturated columns.

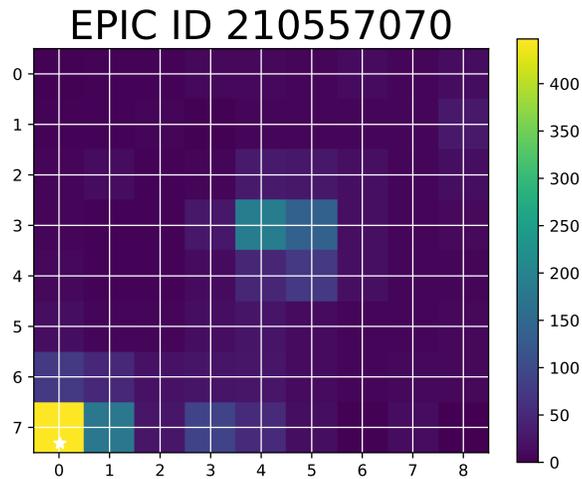


Figure 3.8: The figure above shows an observation where the target of interest is too close to the edge of the postage stamp for the entire PSF to be adequately captured.

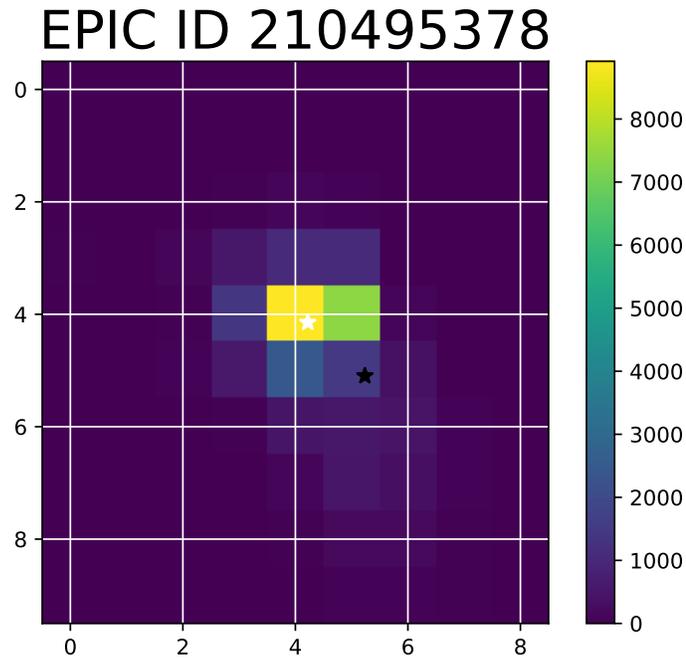


Figure 3.9: The figure above shows a single star observation with the location of the star according to *Gaia* marked with a white star and the location of the star according to the file header marked with a black star.

3.3 PSF Modeling

The first step in creating a PSF model for *K2* was to download observations of stars from *K2*. I chose Campaigns 1, 7, and 13 for a range of campaigns over the course of the *K2* mission and downloaded all observations for channels 37, 41, 44, 73, 75, and 76. Channels 75 and 76 are near the edge of the detector while channel 73 is on the inner corner of the same square of two modules. Channel 37 is near the middle of the field of view one square left of the center, and 41 and 44 are in the square of modules at the center of the field of view. Towards the center of the field of view the PSF should be more stable while towards the edges of the field of view the PSF should be more distorted. It is important to check the

PSF in both of these locations for this reason. The PSF also varies by channel so channels may not be similar to rotationally symmetric channels. The channels of interest are highlighted in blue in the image in Figure 3.10.

To begin stacking, the observations were sorted by channel so that only observations in the same channel were stacked. This was done to account for any differences in PSF shape due to location within the *K2* field of view, since the PRFs for *Kepler* are known to vary by channel.

The first attempt at aligning the arrays was to pad each array with zeros in order to shift the center of the image towards the center of the array. This made the arrays the same size so that they could be added, but the centers were not close enough. The second attempt at aligning involved inserting each array into an empty, larger array and aligning the center of each observation array with the center of the large array. This more closely aligned the centers, but they still could not be aligned at sub-pixel resolution.

In order to better stack the PSFs, the images had to be re-sampled in order to split the available flux data between more pixels to increase resolution and more closely align the stellar centers. This was done using the `scipy` package `ndimage.interpolation.zoom`. Prior to interpolation, the arrays had to be checked to make sure no negative values were present, as negative values created NAN (not a number) values when interpolated. This was done by determining the smallest positive value in the array and replace all negative values with it. Each observation was interpolated by a factor of 3, increasing the pixel resolution of each observation by a factor of three. Doing this allowed the centers of the stars in each image to be more closely aligned.

Kepler Focal Plane Layout

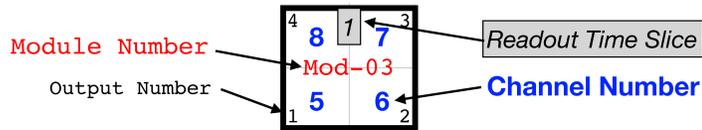
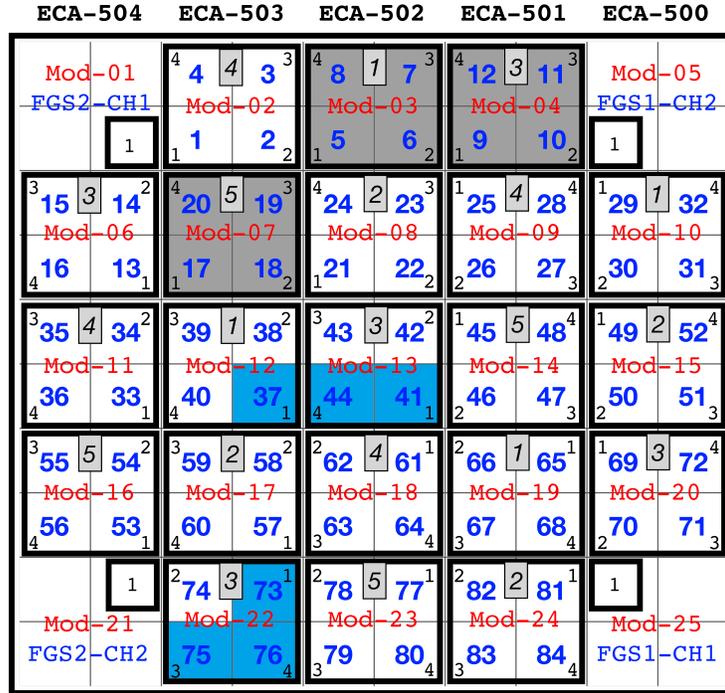


Figure 3.10: K2’s focal plane showing each module, channel, and readout number. The greyed out modules were not operational for the full duration of the mission. The blue modules are the channels being evaluated for PSF in this paper. Image credit: Johnson (2015), edited to highlight channels of interest

3.3.1 PSF Orientation

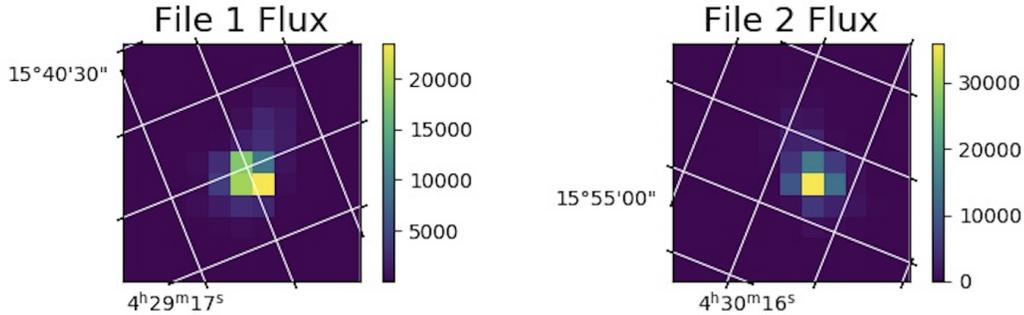


Figure 3.11: Diagram showing two observations from Campaign 13 with different WCS orientation.

Each *K2* observation, even within the same campaign, has a different set of WCS coordinates. Differing orientations could potentially pose a problem to stacking if the orientations matter in the PSF shape. The PRF models for *Kepler* are not symmetric.

However, the PSF models are being constructed by channel and by campaign. Since the space craft is pointed towards the same field of view for each campaign, and the PSF models are being constructed for each channel and campaign uniquely, the rotational effects on PSF shape should not effect the PSF models made through the approach highlighted above.

3.4 Summary of Approach to *K2* PSF Creation

The PSF stacks will be constructed for each channel in each campaign, rather than for a channel across a campaign or a campaign across a channel. There are factors that change PSF shape between channels, such as distance from the center of the field of view. There are also factors that change PSF shape between

campaigns, such as the field of view, the stars in that view, and the instrumental effects present during that portion of the mission. The observations used to create these stacks must also be evaluated to make sure they are single observations, as the presence of multiple stars, saturated stars, or stars without the PSF fully captured would not contribute an accurate PSF shape to the model. The next chapter will show the outcome of the PSF stacking process described in this chapter, and discuss the implications of the results for *K2* PSF modeling.

Chapter 4

Results

I sorted 4022 observations from Campaigns 1, 7, and 13 in channels 37, 41, 44, 73, 75, and 76 using `astroquery` with the *Gaia* database to determine which files were single and which contained background stars. This technique was able to successfully pick out 1597 observations that contained background stars. However, some observations escaped this filter.

As shown in the previous chapter, there were still multi-star observations, bright stars, cut off observations, and other forms of contamination that remained, even after sorting the results with *Gaia*. These remaining contaminated observations were removed by individually inspecting a flux still for each observation. A total of 421 additional observations were eliminated this way.

Once these two processes were complete, a total of 2004 uncontaminated observations remained. Of these remaining observations, almost all were from Campaigns 1 and 13. Campaign 7 resulted in only 50 usable, uncontaminated files between the six channels. With such a small remaining sample, Campaign 7 was determined to be unusable for PSF modeling.

4.1 Single Star Observations By Channel and Campaign

The observation totals for each channel and campaign, as well as removals for each channel and campaign at each step are broken down into more detail in Tables 4.1, 4.2, and 4.3. The total observations in each channel for each campaign is given, followed by the number of observations removed through *Gaia* identification of multi-stars observations, followed by the amount of other contaminated observations individually removed by eye, and finally listing the number of remaining uncontaminated, single star observations. The amount of observations for Campaigns 1 and 13 are similar, with with similar amounts of contaminated observations removed at each step of the process. Campaign 7, comparatively, started out with less observations and the majority of the observations in each campaign were removed.

The breakdown of total observations for each channel in each campaign and the amount of observations removed at each step is also given in percentages in Tables 4.4, 4.5, and 4.6. For Campaigns 1 and 13, around a quarter of the observations were usually multi-stars detected by *Gaia*, less than 10 percent of the observations were otherwise contaminated and individually removed (with the exception channel 76 in Campaign 13), and more than half of the observations were uncontaminated, single stars (again, with the exception of channel 76 in Campaign 13). However, on average channel 7 had more than three quarters of the observations determined to be multi-star observations by *Gaia*, slightly more than 10 percent of the observations contaminated but not excluded by *Gaia*, and less than 10 percent of the observations as usable, non-contaminated single stars. It

is worth noting that the remaining, single star observations were heavily weighted towards the inner channels, with between 12.7 and 19 percent of the observations making it through the channel vetting, compared to less than 1 percent of the observations for the outer channels.

For the typical *K2* campaign, more than half of the observations are uncontaminated, single star observations. This assumes Campaigns 1 and 13 can be considered typical since there is so much variation between campaigns. Campaign 7, however, only has around 15 percent of the observations as uncontaminated for the inner channels, and less than 1 percent of the observations for the outer channels. More work must be done to determine how typical the percentages in Campaign 1 and 13 are for the rest of the *K2* missions.

Observations Per Channel (Campaign 1)

Channel	Observations	<i>Gaia</i> Multi-Star	Other Contamination	Single Star
37	258	74	13	171
41	265	68	9	188
44	361	75	88	198
73	253	65	25	163
75	279	94	19	166
76	286	67	21	198
Total	1702	443	175	1084

Table 4.1: Summary of files for Campaign 1. The first column gives the total number of files in each channel, followed by the number of files removed via Gaia query as multi-stars, followed by the number of files removed via visual inspection as multi-stars or bright stars, followed by the remaining unsaturated, single stars.

Observations Per Channel (Campaign 7)

Channel	Total Observations	<i>Gaia</i> Multi-Star	Other Contamination	Single Star
37	63	37	14	12
41	118	89	14	15
44	116	88	7	21
73	138	122	15	1
75	216	185	30	1
76	160	149	11	0
Total	811	670	91	50

Table 4.2: Summary of files for Campaign 7. By the final step of this process for this campaign, too few files were left to create an accurate PSF model.

Observations Per Channel (Campaign 13)

Channel	Total Observations	<i>Gaia</i> Multi-Star	Other Contamination	Single Star
37	285	104	11	170
41	341	126	19	196
44	376	100	12	264
73	178	55	14	109
75	156	39	12	105
76	173	60	87	26
Total	1509	484	155	870

Table 4.3: Summary of files for Campaign 13. Numbers of observations at each step are closer to Campaign 1 than to Campaign 7.

Observations Per Channel (Campaign 1)

Channel	<i>Gaia</i> Multi-Star (%)	Other Contamination (%)	Single Star (%)
37	28.7	5.04	66.3
41	25.7	3.40	70.9
44	20.8	24.4	54.8
73	25.7	9.88	64.4
75	33.7	6.81	59.5
76	23.4	7.34	69.2
Average	26.3	9.48	64.2

Table 4.4: Summary of files for Campaign 1 by percentage.

Observations Per Channel (Campaign 7)

Channel	<i>Gaia</i> Multi-Star (%)	Other Contamination (%)	Single Star (%)
37	58.7	22.2	19.0
41	75.4	11.9	12.7
44	75.9	6.03	18.1
73	88.4	10.9	0.72
75	85.6	13.9	0.46
76	93.1	6.88	0.00
Average	79.5	12.0	8.55

Table 4.5: Summary of files for Campaign 7 by percentage.

Observations Per Channel (Campaign 13)			
Channel	<i>Gaia</i> Multi-Star (%)	Other Contamination (%)	Single Star (%)
37	36.5	3.86	59.6
41	37.0	5.57	57.5
44	26.6	3.19	70.2
73	30.9	7.87	61.2
75	25.0	7.69	67.3
76	34.7	50.3	15.0
Average	31.8	13.08	55.1

Table 4.6: Summary of files for Campaign 13 by percentage.

4.2 PSF Stacks by Channel

Once the observations were filtered to only usable, uncontaminated single stars, these files were then stacked to form an average PSF model. This was done by adding the flux from each observation onto a larger empty array, being sure to align the centers of the stars. Once the observations were all accumulated in this way, the larger conglomerate of flux was trimmed of any columns or rows containing only zeros, and normalized by dividing the entire array by the largest flux value so that the model ranged from a highest value of 1 to a lowest value of zero.

The results are shown in Figures 4.1 and 4.2 (Campaign 1), Figures 4.3 and 4.4 (Campaign 7), and Figures 4.5 and 4.6 (Campaign 13). The PSF stacks are shown both with only observations determined to be multi-star by *Gaia* and with additional contaminated results removed. The details for each campaign are discussed in more detail below.

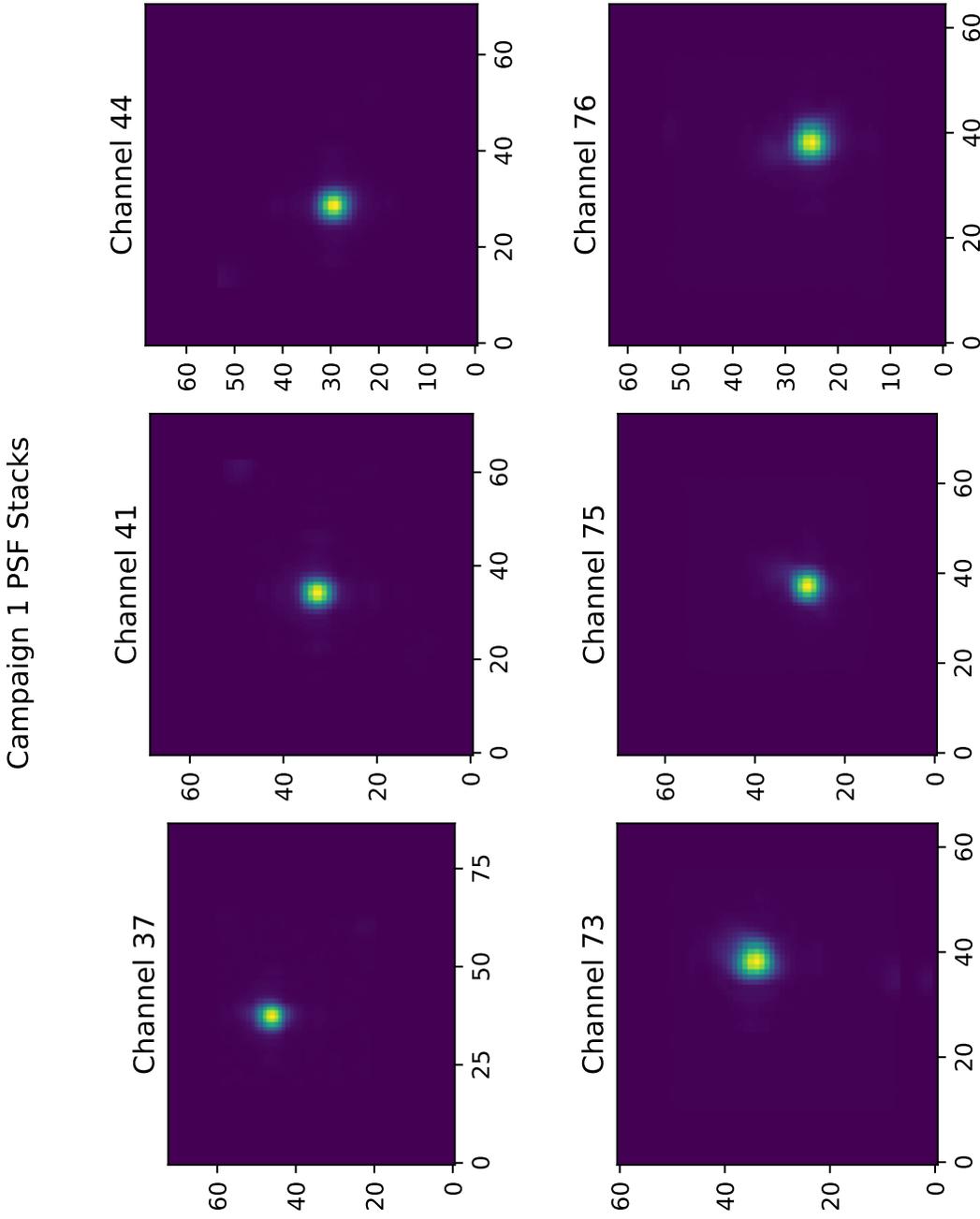


Figure 4.1: This figure shows the PSF stack of observations for each channel in Campaign 1, after trimming of empty columns and rows.

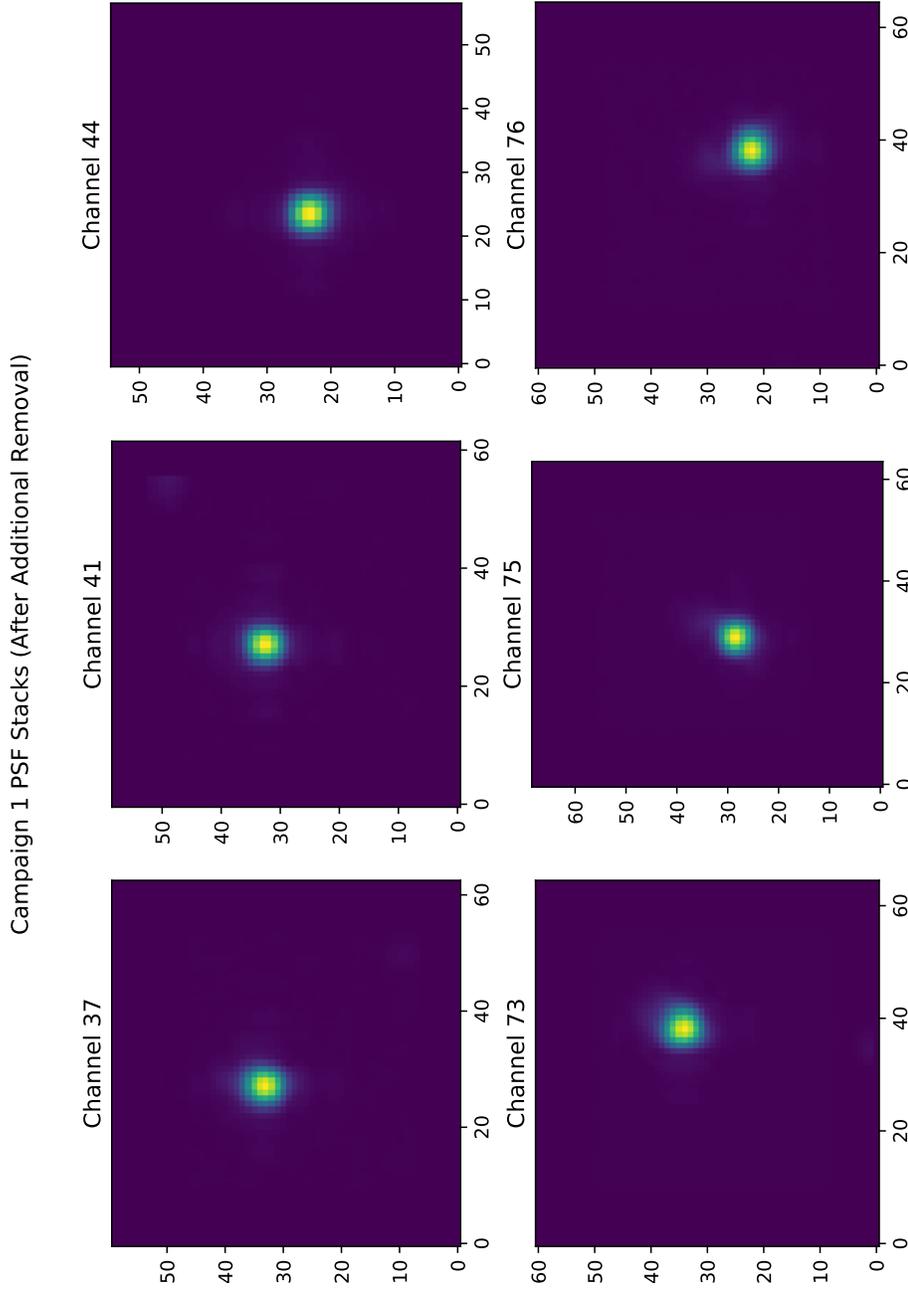


Figure 4.2: This figure shows the PSF stack of observations for each channel in Campaign 1, after additional contaminated files were individually removed. For this campaign, the difference before and after additional removal is minimal.

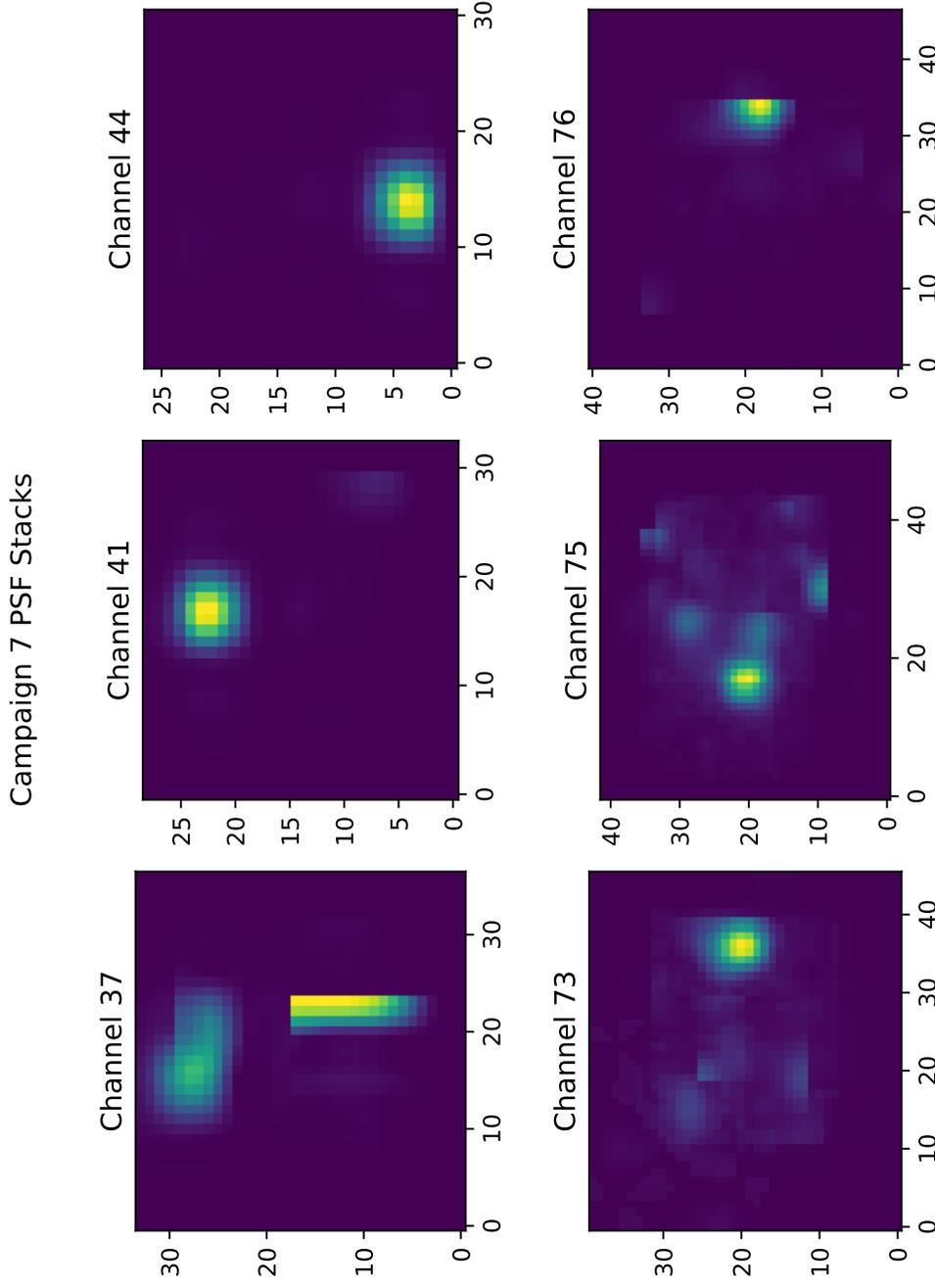


Figure 4.3: This figure shows the PSF stack of observations for each channel in Campaign 7, after trimming of empty columns and rows. The observations are very contaminated.

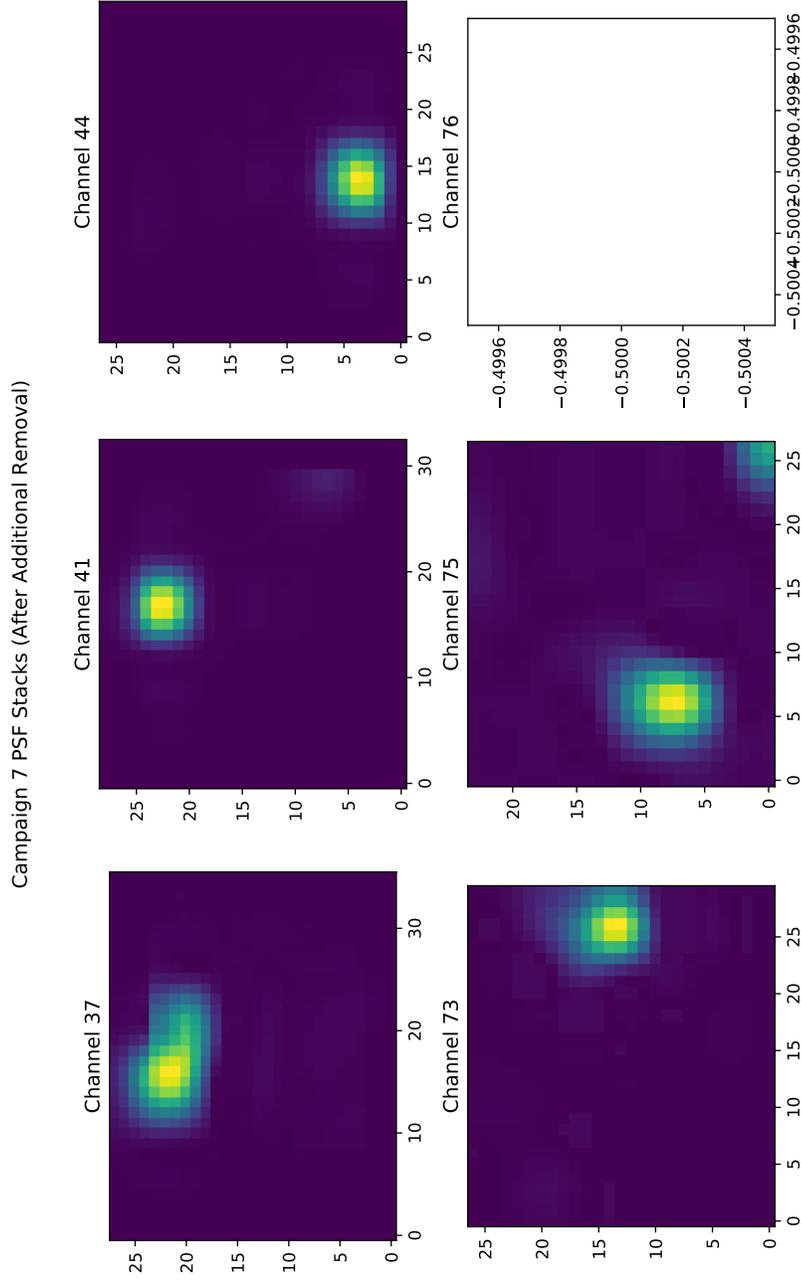


Figure 4.4: This figure shows the PSF stack of observations for each channel in Campaign 7, after additional contaminated files were individually removed. Even removing as many obviously contaminated observations as possible, the PSF stacks for Campaign 7 are still poor quality. The bottom right image is empty, since all the observations in channel 76 were removed.

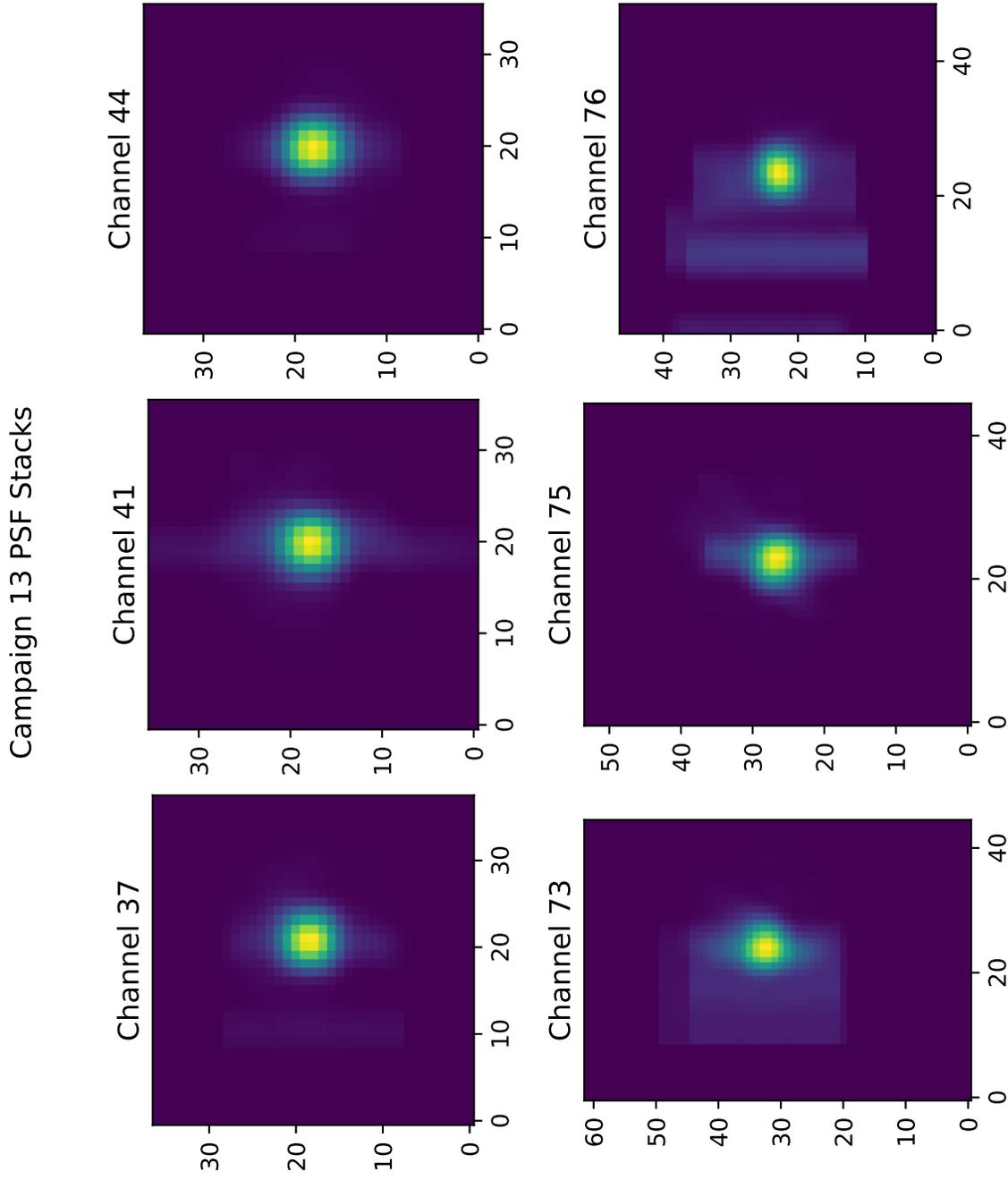


Figure 4.5: The figure above shows the stacked observations for each channel in Campaign 13, prior to additional visual inspection and removal of contaminated observations. This campaign has visible saturation from bright stars.

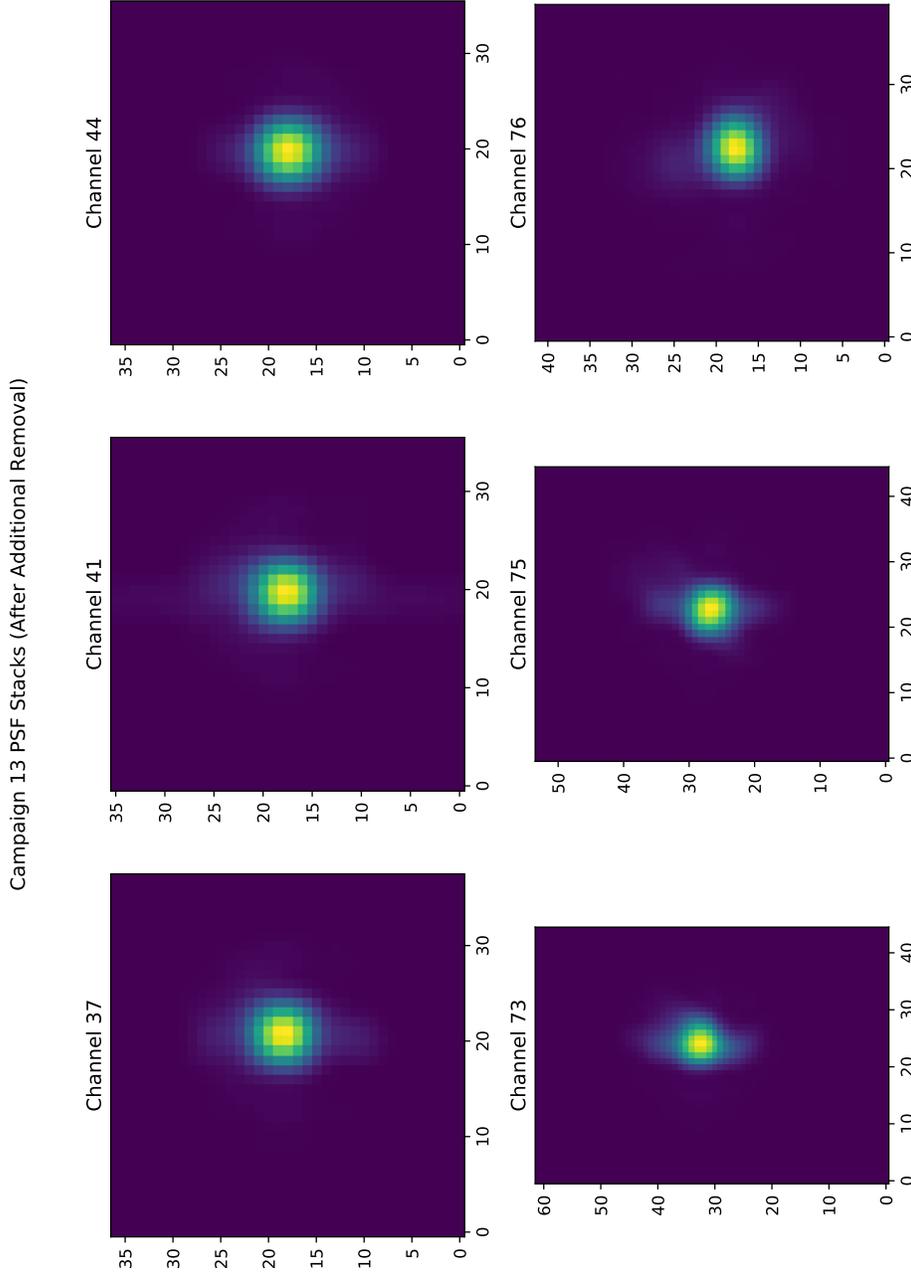


Figure 4.6: This figure shows the PSF stack of observations for each channel in Campaign 7, after additional contaminated files were individually removed. Most of the saturation from bright stars has been removed though there is still some residuals that look like bleeding. It may be prudent to re-examine the data by eye and with a stricter threshold for bright stars.

4.3 Discussion of Initial Results

4.3.1 Campaign 1

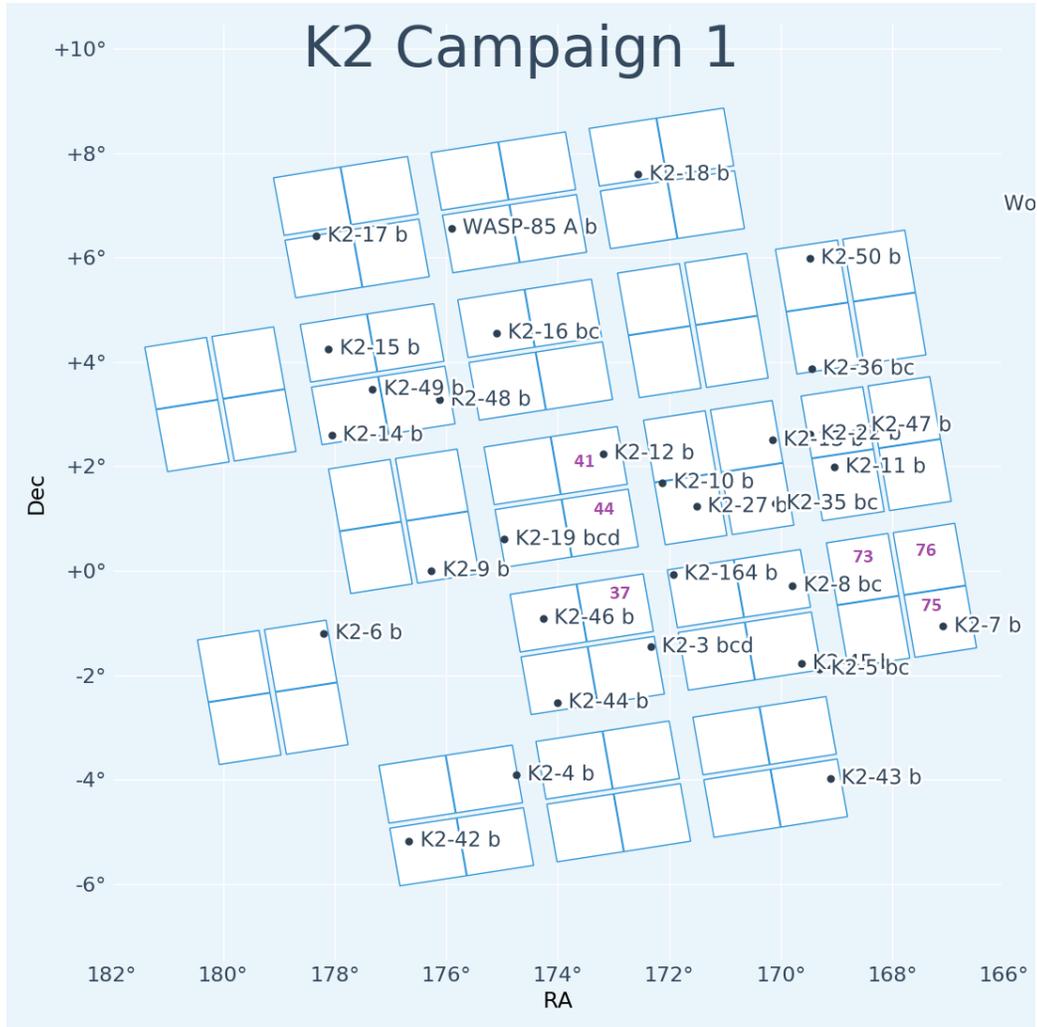


Figure 4.7: Figure showing the field of view for *K2* during Campaign 1. Channels of interest are labeled in purple. Image credit: <https://keplerscience.arc.nasa.gov/k2-data-release-notes.html#k2-campaign-1>

Campaign 1 resulted in the most uniform PSF stacks. There are no areas of particularly high stellar density for this Campaign and the majority of the stars were single,

uncontaminated observations of each channel, resulting in sufficient observations for PSF stacking.

4.3.2 Campaign 7

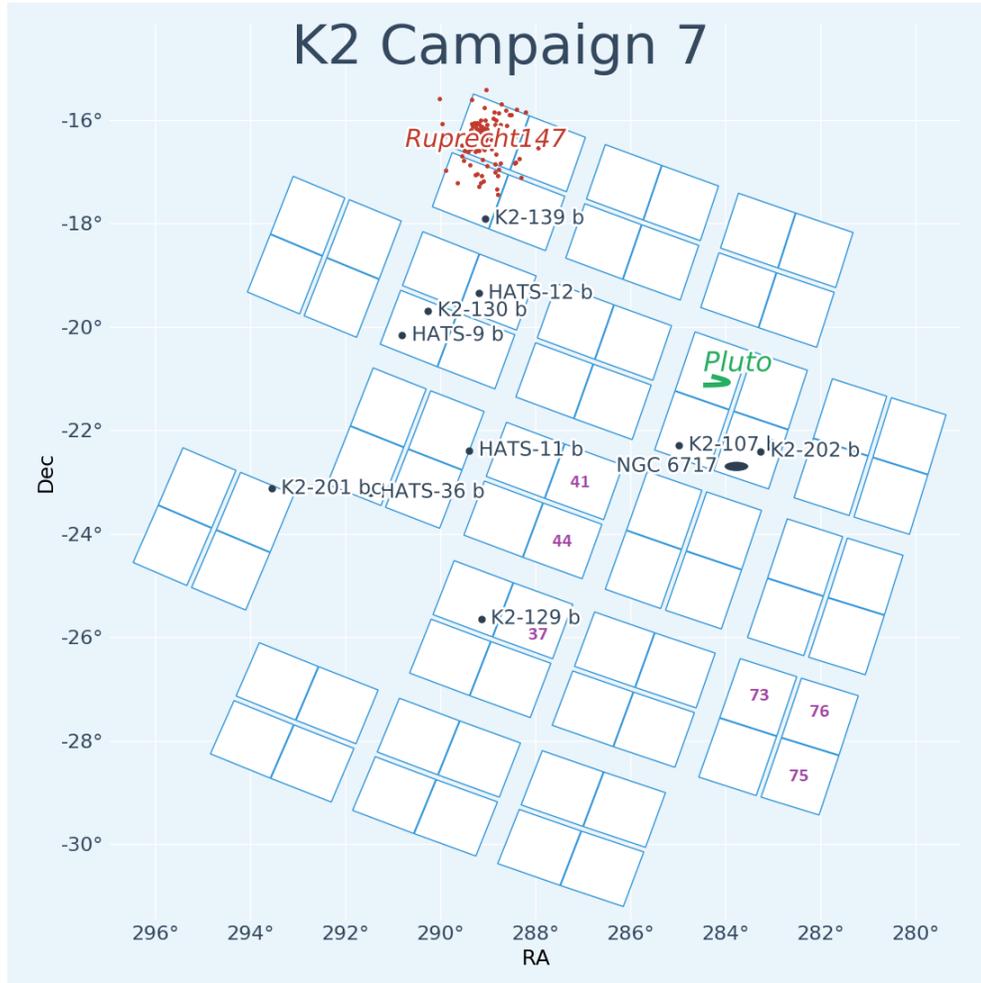


Figure 4.8: Figure showing the field of view for *K2* during Campaign 7. Channels of interest are labeled in purple. Image credit: <https://keplerscience.arc.nasa.gov/k2-data-release-notes.html#k2-campaign-7>

Campaign 7 ended up being a poor candidate for PSF modeling. This campaign had a lower number of files than the other two campaigns investigated, only about half the amount. Campaign 7 had a lower number of observations returned than previous *K2*

campaigns due to a worse compression estimate than was used for previous campaigns (Colon 2016). There was a higher stellar density than previous campaigns, leading to more pixels being required to store observations and the data not being able to be as easily condensed. Since on-board data space is limited and the observations were larger, less targets could be stored.

Additionally, Campaign 7 used an alternate, lower gain antenna for collection, resulting in an increase in roll drift (Colon 2016). Roll drift is one of the biggest influencing factors in reducing precision for *K2*. A lower gain antenna means that the beam is wider, making it easier to capture observations with worse precision. In combination with the field being particularly crowded, this low gain antenna resulted in particularly poor photometry compared to other campaigns.

Campaign 7 ended up with no single, uncontaminated files for channel 76, and only one file that was single and uncontaminated for channel 75 and 74. The channels nearer the edge of the field of view, where the PSF was less defined for *Kepler* had less usable files than the inner channels. From this point on, Campaign 7 will no longer be discussed in terms of its PSF, since the data is not viable.

4.3.3 Campaign 13

Campaign 13 had much more usable data than Campaign 7, but had more saturated, bright stars than Campaign 1. The majority of these stars were removed during individual inspection, but there is still some traces of saturation even in the final PSF stack. It is difficult to tell what stars are saturated versus what stars are bright but unsaturated at the more intermediate region.

Channel 76 in Campaign 13 was also unique in the amount of contaminated targets removed. All the channels besides channel 76 had more than half of the files as uncontaminated, single star observations. However, channel 76 had only 15 percent of observations that were uncontaminated, single stars. Looking at the field of view for

Campaign 13 in Figure 4.9, the reason for this becomes clearer. Channels 73, 76, and 75 are all viewing the Hyades open cluster, with channel 76 being the closest to the center of it. For this reason, channel 76 has a disproportionate number of multi-star observations.

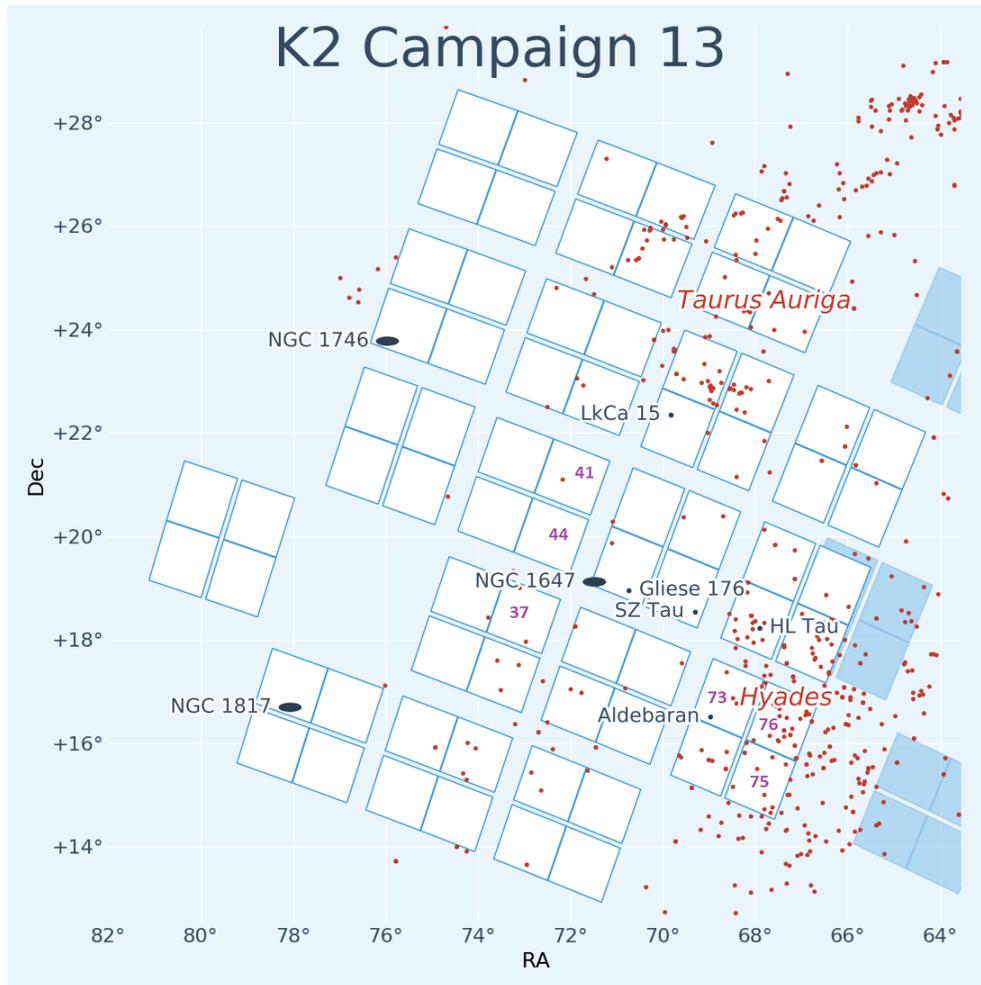


Figure 4.9: Figure showing the field of view for *K2* during Campaign 13. Channels of interest are labeled in purple. Image credit: <https://keplerscience.arc.nasa.gov/k2-campaign-13-science-program-now-available.html>

4.4 Evaluating the PSF Model

Once the PSF stacks are constructed from aligning and combining all the images, this final amalgamation of flux can be compared to the other two PSF model options: the *Kepler* PRF and the 2D Gaussian. To do this quantitatively, residuals must be calculated and plotted between the PSF stacks and the other two models, in order to see if either of the two would be a close proxy for the PSF stack model, or if the PSF stack contains unique features not yet accounted for in either of these two models.

Before calculating residuals, the PSF stacks must be normalized. In this case, “normalized” means changing the range of values to values between 1 and 0, with 1 being the maximum value in the array. This was done by first checking if the PSF stack contained any negative values and replacing them with the smallest positive value. Once this was complete, the PSF stack could be normalized by identifying the maximum value in the array and dividing by it. This was done for both the *K2* PSF stack and the *Kepler* PRF models. The Gaussian could be constructed intentionally with values between 1 and 0.

Once the models were normalized they could be subtracted. This was done similarly to the stacking of observations. Each PSF observation stack for each channel in each campaign was subtracted from the *Kepler* PRF for the same channel. This was done by aligning the centers of the observation stack model and PRF model and finding the difference between them. The “center” of each model was defined as the pixel with the highest flux, or due to the normalization, the pixel with the value of “1”. This was done by finding the index of the pixel with the maximum value in each model, and assigning those coordinates as the center for the model. In order to properly subtract, one of the files had to be a large enough array that the values of the array being subtracted would never be subtracted from a value outside of the first array, since the arrays are all different sizes. This was done by inserting the PSF into a large empty array, and

then subtracting the PSF.

The same process when performed for a 2D Gaussian would be a little less extensive. The 2D Gaussian would have its center at the center of the model. Therefore, a center pixel would not have to be selected by max flux and could just be assigned as the center pixel.

4.4.1 PSF vs. PRF

The PSF for each campaign and the *Kepler* PRF are compared for each channel. The inner channels (channel 37, 41, and 44) are grouped together, as are the outer channels (channel 73, 75, and 76). The outer channels should be less uniform in shape due to less precision near the edges of the field of view, and this does seem to be the case comparing the figures of the models for the inner channels (Figure 4.10) versus the outer channels (Figure 4.11).

In addition to the visual comparison, each *Kepler* PRF model was also subtracted from each PSF stack to check how similar the two models were. The results for Campaign 1 are shown in Figure 4.12 (the inner channels) and Figure 4.13 (the outer channels), and for Campaign 13 are shown in Figure 4.14 (the inner channels) and Figure 4.15 (the outer channels).

The subtraction is not exact. Interpolation to the same grid was not successfully completed. The two models were interpolated to increase the resolution by a factor of 3, allowing the centers to be aligned to a sub-pixel level. However, for both the PSF stacking and PSF and PRF subtraction, the exact same scaling is not being used. Due to this, the models are not the exact same size, which makes it difficult to truly compare the shapes. However, even with the imperfect subtraction, it does seem clear that the PSF stacks and *Kepler* PRF models are not identical.

Not only do the PSF stacks for each channel in each campaign not match the *Kepler* PRF models, but the PSF stacks for each channel do not match each other, across

campaign. More stacking with proper interpolation is needed to exactly identify the difference in shape, but it is clear that the PSF stacks for each channel between Campaigns 1 and 13 are not the same. Without a proper third campaign to compare with, it is difficult to say how much the PSF for each channel varies between campaigns. It is possible that all of the campaigns have very different PSF models for the same channel, but it is also possible there are some similarities. More PSF stacking in more channels and campaigns is required.

4.4.2 PSF vs. Gaussian

There was insufficient time to evaluate how well the PSF stacks compared to a 2D Gaussian. However, when visually inspecting the PSF stacks for each channel, the shape for each PSF does not appear to be symmetrical. A 2D Gaussian can be elongated like an oval or more circular but it requires the shape to be symmetrical. The *Kepler* PRF models were not perfect 2D Gaussians. It does not appear that the *K2* PSF models could be properly fit by a 2D Gaussian either.

4.4.3 Comparing PSF between Campaigns by Channel

In addition to comparing the PSF stack for each channel to the *Kepler* PRF in the same channel, the PSF stack for each channel was also compared to the PSF stack in the same channel in the other campaign. The results of the subtraction are shown in Figure 4.16 (the inner channels) and Figure 4.17 (the outer channels).

In addition to the *K2* PSF stack model having significant residuals from the *Kepler* PRF model, the *K2* PSF stacks models in each campaign but the same channel also have significant residuals. The *K2* PSF model clearly differs not only from the *Kepler* PRF for each channel, but also differs within the same channel, across campaign. This confirms that it was prudent to create PSF model for each channel within each campaign, rather than making an average PSF model for each channel across all campaigns.

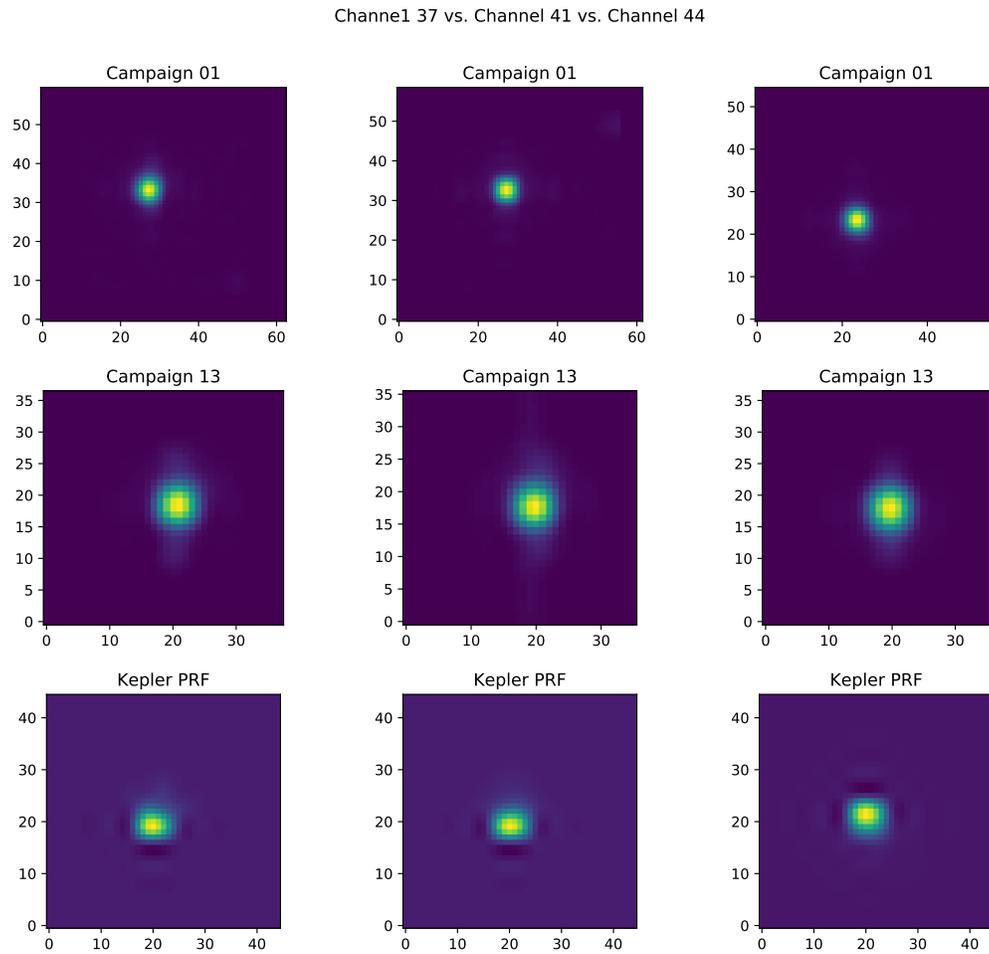


Figure 4.10: The figure above shows the normalized *K2* PSF stack for the inner channels in Campaign 1, Campaign 13, and the normalized *Kepler* PSF for Channel 37(left), Channel 41(middle), and Channel 44(right). These channels are close to the center of the field of view.

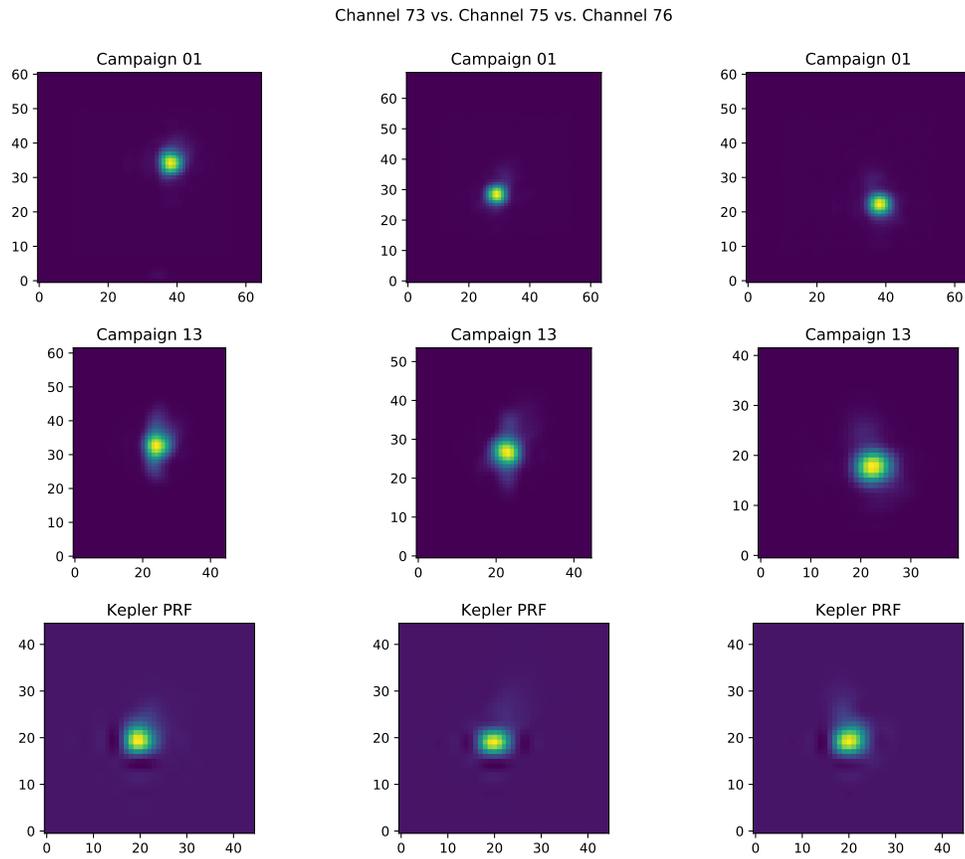


Figure 4.11: The figure above shows the normalized *K2* PSF stack for the outer channels in Campaign 1, Campaign 13, and the normalized *Kepler* PSF for Channel 73 (left), Channel 75 (middle), and Channel 76 (right). These channels are at the edge of the field of view.

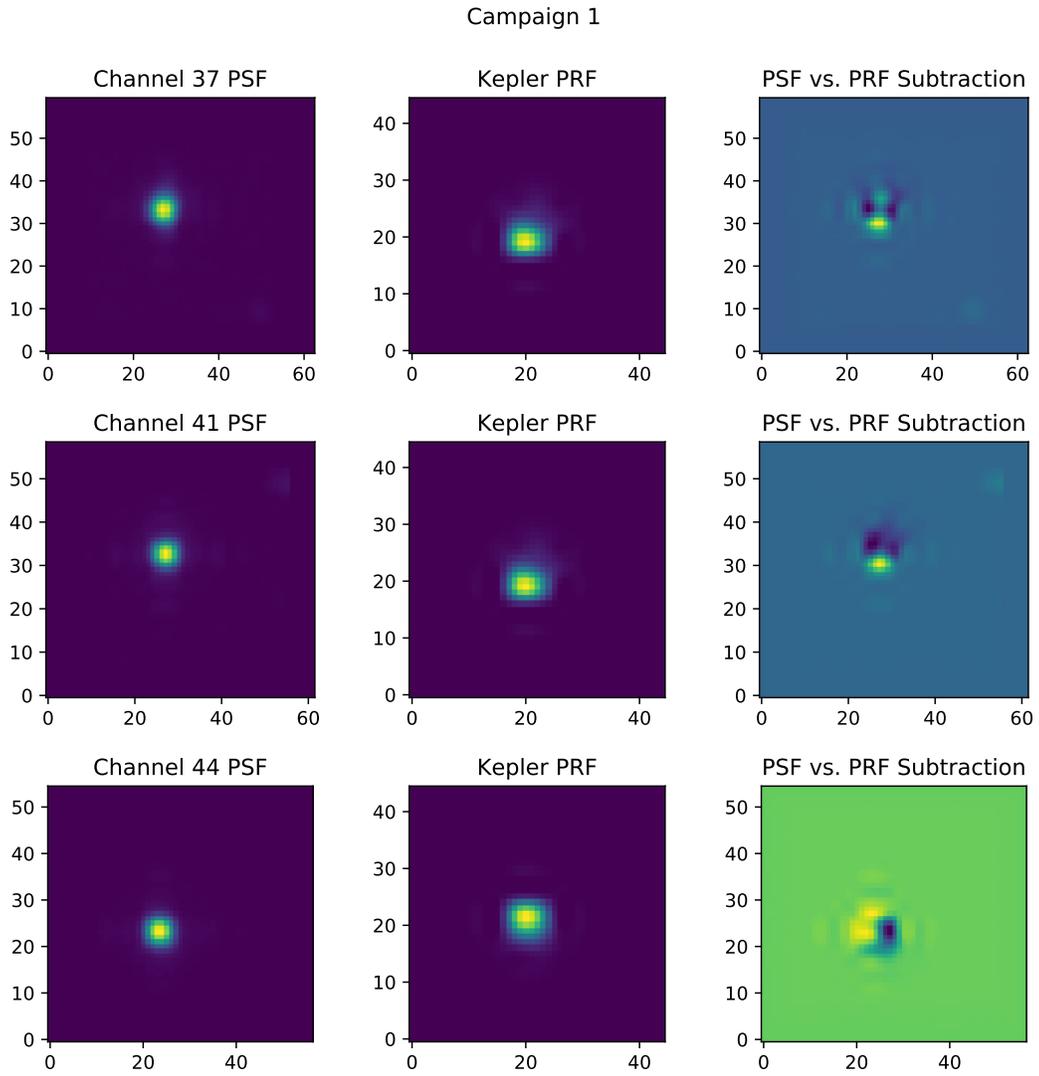


Figure 4.12: The figure above shows the subtraction between the K2 PSF stack (left column) for each of the inner channels (37, 41, and 44) and the *Kepler* PRF model (middle column) for Campaign 1. The residuals are shown in the right column.

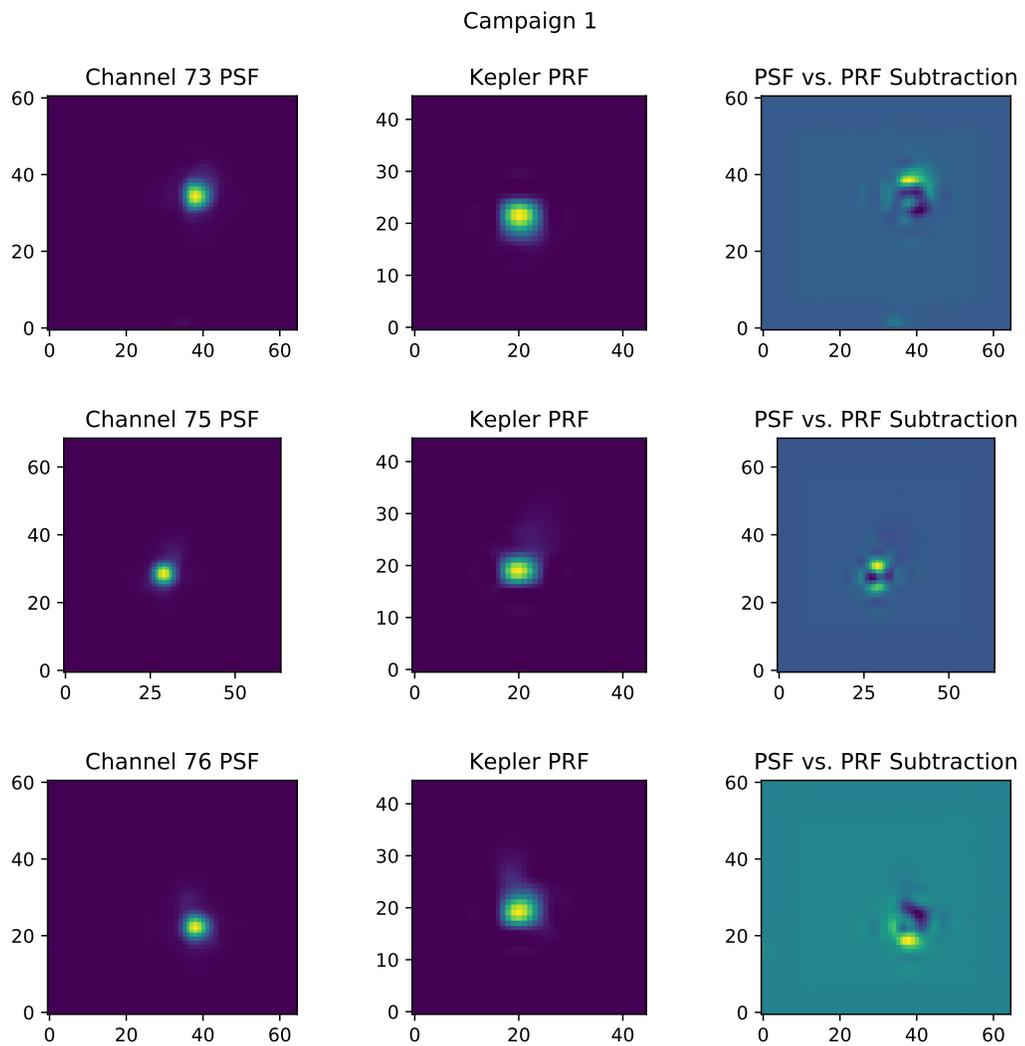


Figure 4.13: The figure above shows the subtraction between the K2 PSF stack (left column) for each of the outer channels (73, 75, and 76) and the *Kepler* PRF model (middle column) for Campaign 1. The residuals are shown in the right column.

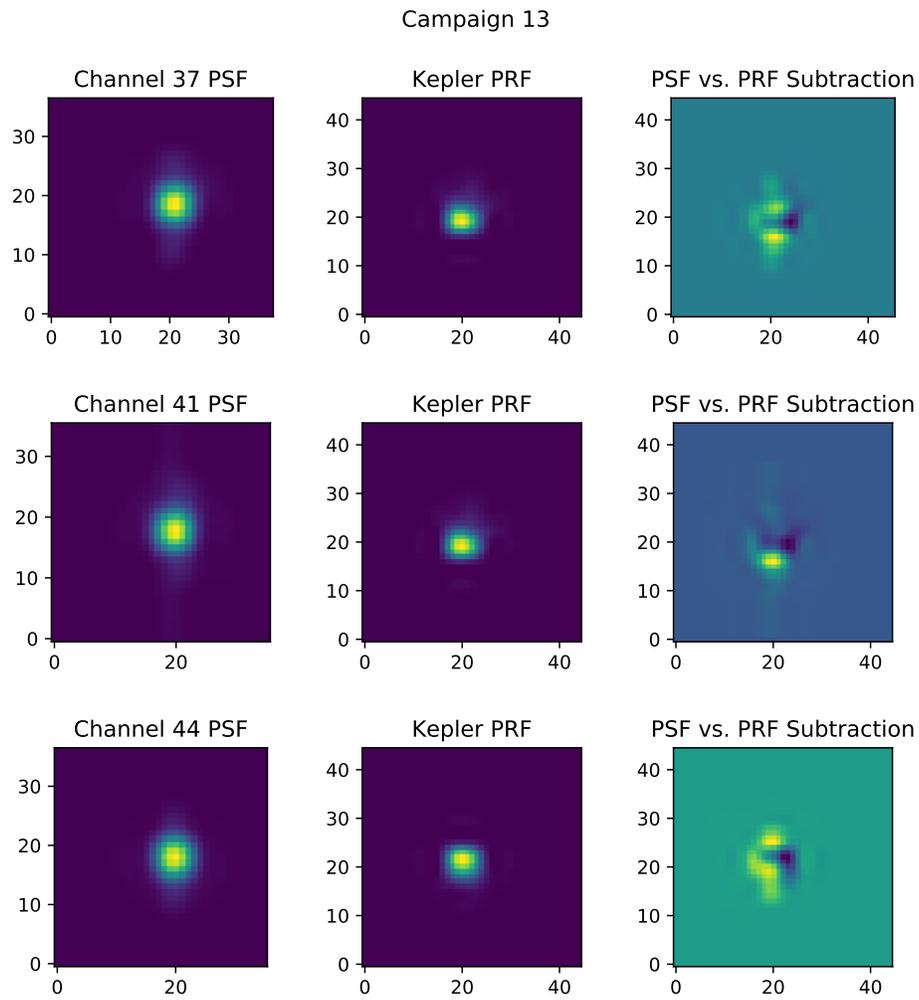


Figure 4.14: The figure above shows the subtraction between the K2 PSF stack (left column) for each of the inner channels (37, 41, and 44) and the *Kepler* PRF model (middle column) for Campaign 13. The residuals are shown in the right column.

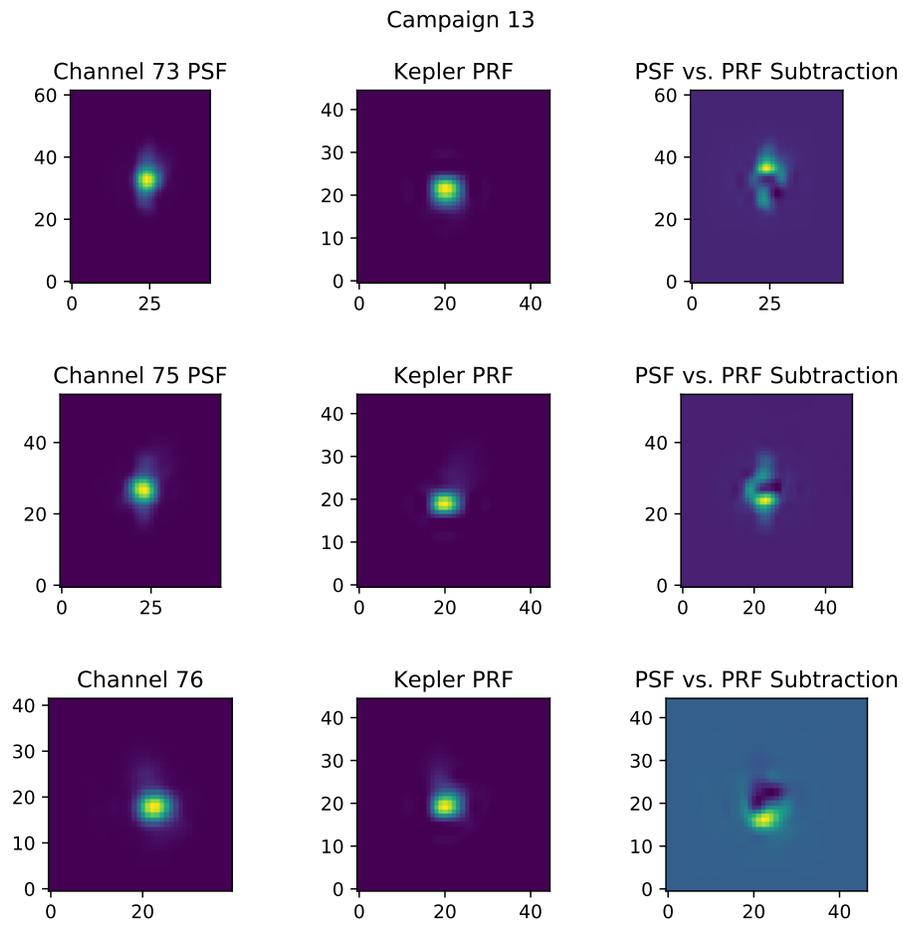


Figure 4.15: The figure above shows the subtraction between the K2 PSF stack (left column) for each of the outer channels (73, 75, and 76) and the *Kepler* PRF model (middle column) for Campaign 13. The residuals are shown in the right column.

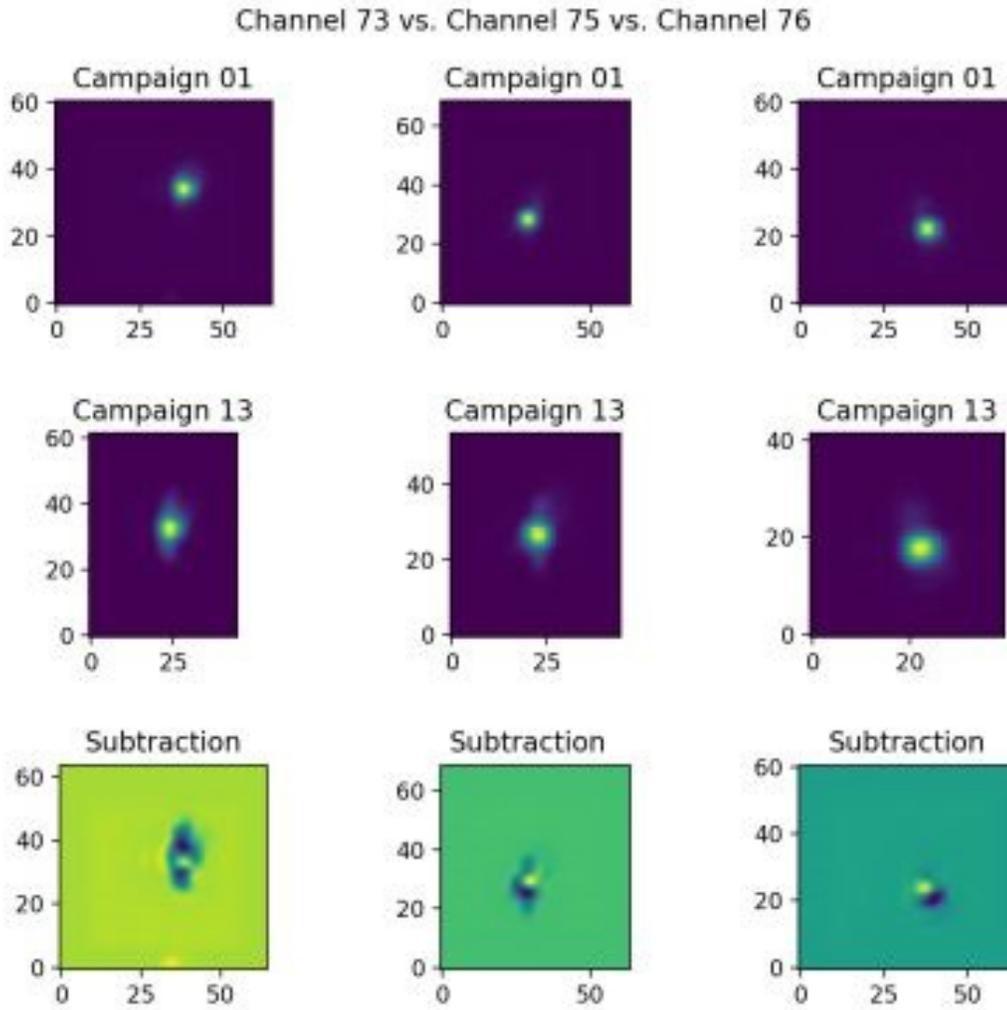


Figure 4.17: The figure above shows the subtraction between the *K2* PSF stack for Campaign 1 (top row) for each of the outer channels (73, 75, and 76) and the *K2* PSF stack for Campaign 13 (middle row). The residuals between the two are shown in the bottom row.

4.5 Future Steps

Even though interpolation made it easier to stack observations, the observations were not able to be fully interpolated to the same scale. Interpolation can be used to increase or decrease resolution by reading out the existing data in more or less pixels. The smaller the pixels, the closer the images can align since the pixel location for the center can be more exact.

However, not all observations in *K2* are the same size, both in terms of pixels and area on the sky. I was unable to interpolate observations to the same scale and thus even with smaller resolution, the centers could not be perfectly aligned. Some of the observations may also be disproportionately large in size due to lack of equal scaling, which might cause widening of the PSF. Some of the observations may also be disproportionately small in size, which might lead to their addition to the PSF being overly concentrated towards the center of the PSF and not reaching the edges.

The same is true for the subtraction. The PSF stacks, although normalized, were not exactly at the same scale as the Kepler PRFs. This was helped somewhat by interpolating, in terms of increasing pixel resolution, but was not fully corrected. The scale of both models was similar but not identical. The models were both normalized in terms of flux, but were not the same in terms of spatial resolution.

The PSF does appear to change with not only each channel, but also between the same channel in different campaigns. Brighter stars require more pixels to capture flux, leading to more pixel space. One factor that was not considered in the PSF stacking was the brightness of the observations, apart from making sure that they did not saturate the CCD and have large bleed columns. In order to more accurately remove background stars, it might be ideal to make multiple PSF stacks per channel in each campaign, for different ranges of stellar magnitude.

Chapter 5

PSF Model Applications

These PSF models can be used to help refine light curves for stars that have not yet been evaluated for exoplanets. This includes bright stars that were typically avoided in the *Kepler* mission but observed during the *K2* mission. The original *Kepler* mission was purposefully set to avoid bright stars, as they were known to saturate the CCD. However, when *Kepler* was reborn as *K2*, the field of view became every-changing along the ecliptic, resulting in a much higher number of bright targets. These bright targets would be excellent candidates for followup studies requiring high flux, such as atmospheric characterization. For this reason, it is important to get accurate light curves for bright stars to look for any possible exoplanet transits.

5.1 Application to *Halophot*

The main goal of creating a PSF model for *K2* was to use this model to subtract background stars from observations of bright stars before using *halophot* to extract the light curve from the pixels in the “halo” region around the bright star.

In a recent paper, the authors of *halophot* extracted light curves for 161 of the brightest stars from *K2* Campaigns 4 to 18 (Pope et al. 2019). The goal was to study these bright stars for stellar variability and pulsations, not for exoplanets. However, light curves are required for both, in order to study the periodicity of stellar flux variations. Each observation was processed using *halophot*. This “raw” light curve was then detrended using the *k2sc* (K2 Systematics Correction) detrending method (Pope et al.

2019).

5.2 Transits Around Pre-processed Bright Stars

I downloaded the 161 light curves that were pre-processed using `halophot`. Using the `BoxLeastSquares` package from `astropy.timeseries`, I created a periodogram for each light curve in order to look for possible planets transits.

A Box Least Squares fit (BLS) is frequently used to search light curves for exoplanet transits. This method operates on the assumption that the light curve has two main levels: high flux (when the planet is not in front of the star) and low flux (when the planet is in front of the star) and spends less time at the lower flux than the higher flux (Kovács et al. 2002). Using this simple model of a transit, varying periods are fit to the flux timeseries. In the last step, a periodogram is created, which determines which period occurs the most frequently within the data. Often with BLS fits, multiples of the actual period will show up as peaks, with less frequency.

An example BLS fit is shown in Figure 5.1. This includes the timeseries, the BLS periodogram, and the folded timeseries at the most frequently appearing period. This example was constructed from simulated, low signal to noise data to demonstrate the potency of the BLS method (Kovács et al. 2002).

Unfortunately, in the 161 bright stars, none of the observations returned any clear periods with the BLS fit. Each observation was plotted with the light curve, detrended light curve, and periodogram resulting from the BLS fit. An example of one case is shown in Figure 5.2. Rather than a clear tall peak followed by several short peaks at multiples of the period, there are peaks interspersed at most period intervals.

The only observations that did pop up with seemingly significant periods were ones like that shown in Figure 5.3. This bright star appears to have a possible transit with a period of 3 days. However, a star bright enough to saturate the CCD would be far too large to have a planet close enough to it for the period to be this short, since brighter

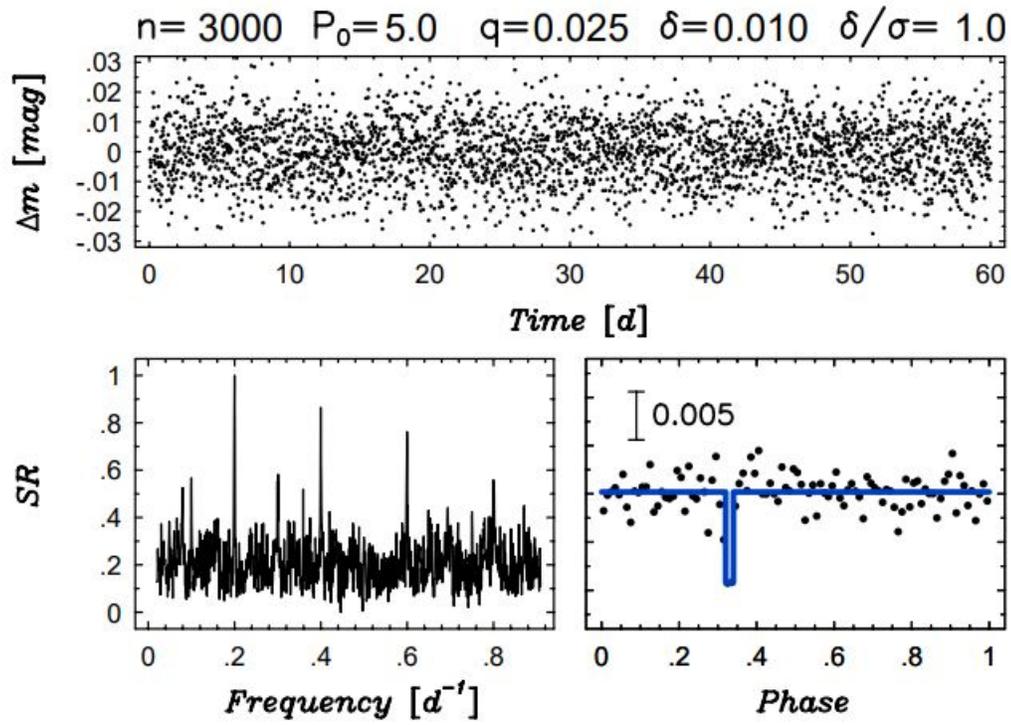
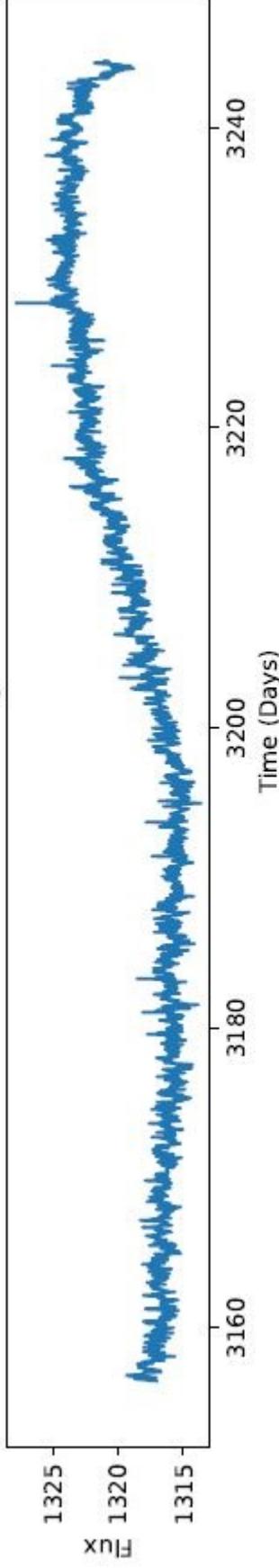


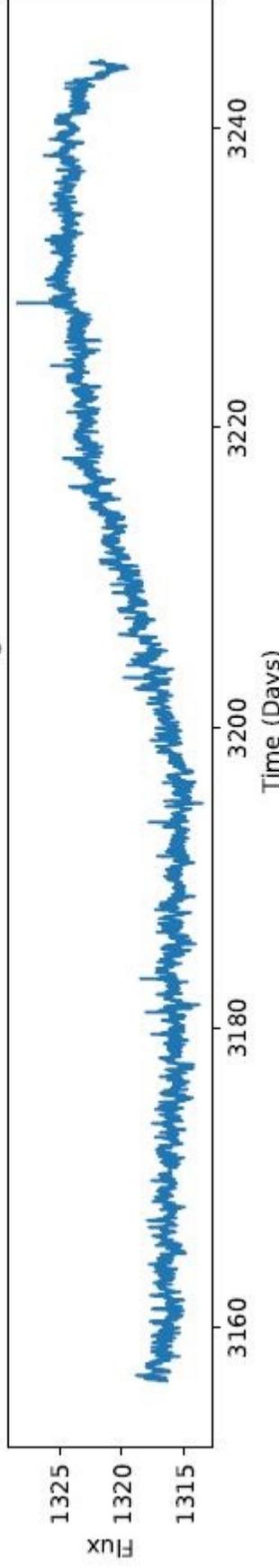
Figure 5.1: The figure above shows a sample Box Least Squares Periodogram (bottom left) extracting a period from a time series with high signal to noise (top), along with the folded time series at the detected period (bottom right). Figure Credit: Kovács et al. (2002)

Periodogram for EPIC ID 200194921

Raw Halo Lightcurve



Detrended Halo Lightcurve



Periodogram

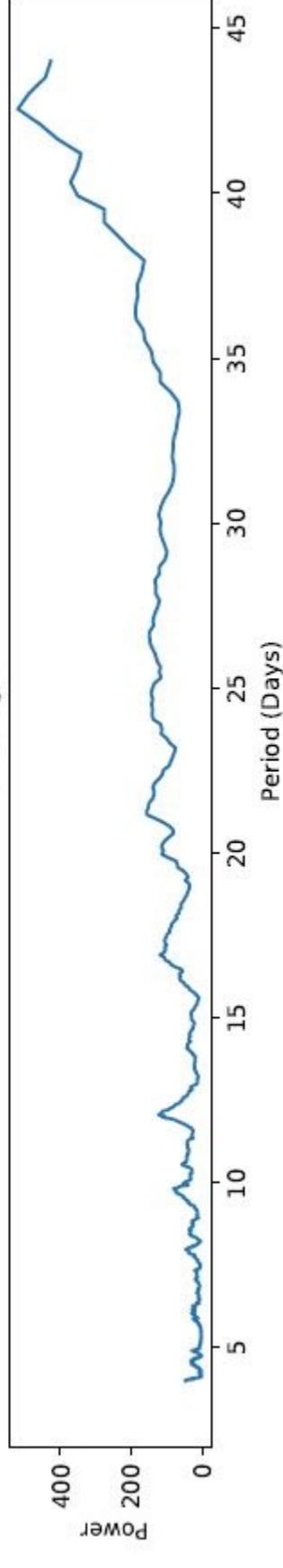


Figure 5.2: The figure above shows the raw light curve from *halophot*, the detrended light curve, and the periodogram for an observation with no remarkable period.

stars are larger and a period this short would require a small semi-major axis. This observation is an example of a false positive, most likely a variable star. Variable stars typically have jagged, almost sinusoidal light curves, as shown in Figure 5.3.

Unfortunately, no clear planet signals were found within this sample of bright stars. However, the first three campaigns of K2 were not included in this sample and may contain planets around them. Furthermore, these light curves were extracted using `halophot` and the `deathstar` protocol for subtracted background stars described above. Using the PSF models developed in this paper to subtract the background stars may result in more refined light curves for these 161 bright stars, and potentially missed transit signals.

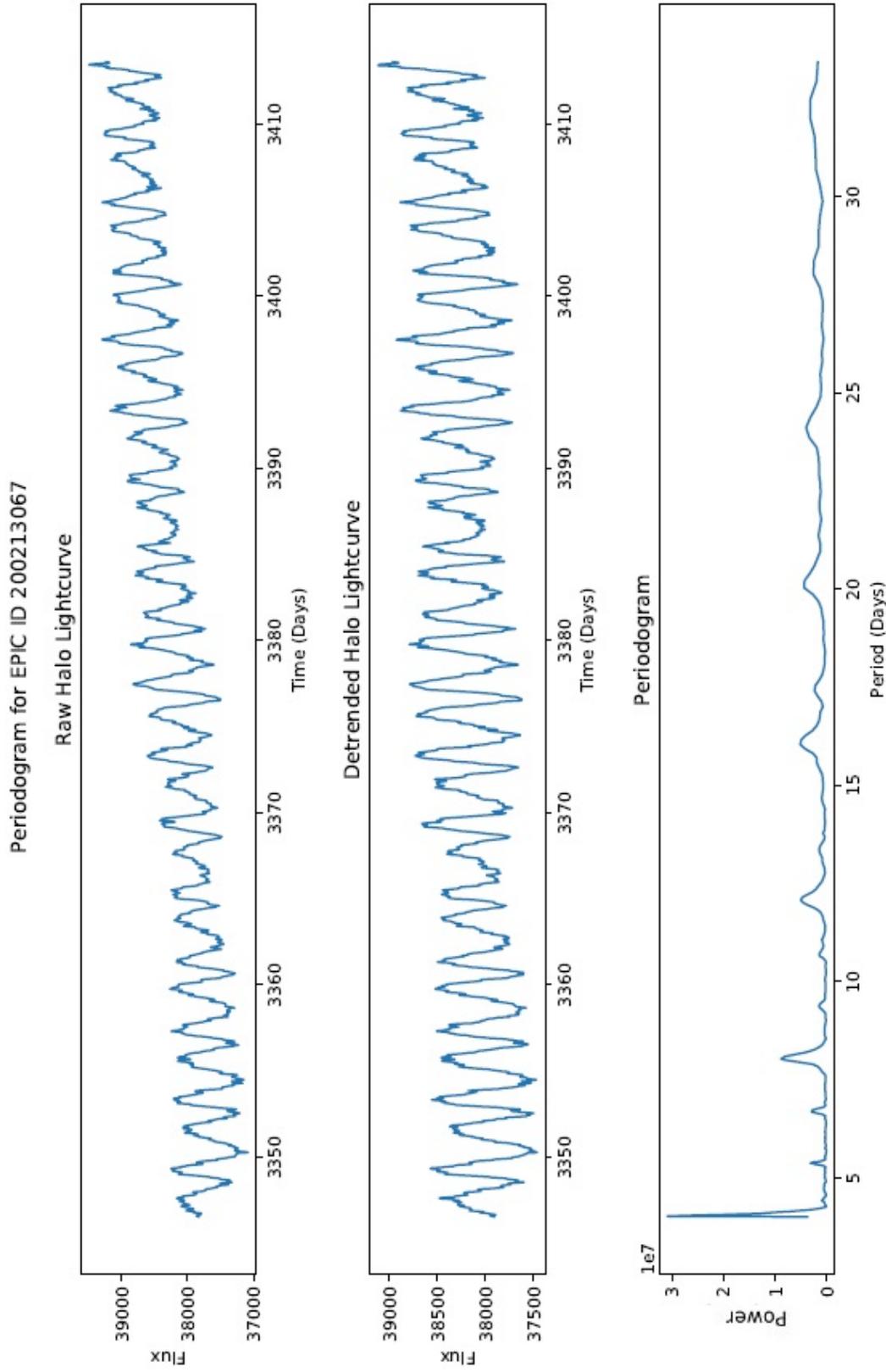


Figure 5.3: The figure above shows a periodogram run for one of the 161 bright targets. This BLS run resulted in a false positive period, indicating a variable star.

Chapter 6

Conclusion

Developing a PSF for *K2* was a complex process. *K2* is subject to more instrumental difficulties than *Kepler*, due to the loss of a reaction wheel, which resulted in not only worse precision due to worse pointing, but also required the spacecraft to be stabilized with thruster events, which caused a zigzag pattern in the data that needed to be detrended and removed. The *K2* mission was also evolving throughout its course, with significant changes between each campaign, both in field of view and instrumental conditions.

Despite all this, *K2* remains an important source of data for exoplanet studies. *Kepler* was able to get longer periods of data for a select region, but *K2* viewed much more of the sky. This allowed the second light of *Kepler* to view many more stars than the original mission was ever intended to, even if the quality was not as good.

6.1 Future Steps for PSF Modeling

This paper begins answering the questions of what a PSF model looks like for a given observation in *K2*. It is clear that the PSF model does not look like a perfect double Gaussian. The PSF models also do not match the *Kepler* PRF models, since they definitely vary between campaigns.

In order to determine how much the PSF models vary by campaign, it is imperative that the process applied above be applied to more campaigns. With Campaign 7 data being too contaminated to properly form a PSF in the channels examined in this project

(channel 37, 41, 44, 73, 75, and 76), there were only PSF models for two campaigns, making it difficult to draw conclusions about how much the PSF models vary by campaign. It is also undetermined how much of these differences would be due to change in time versus change in location, as both are changing each *K2* mission. The effects of time on the PSF could potentially be examined by making a model of PSF in the first half of the campaign and comparing it to a model of the PSF in the second half of the campaign. However, quantity of observations is a concern. Some channels in some campaigns already had too few observations to make an accurate PSF model. Mainly Campaign 7 across the board, but also channel 76 in Campaign 13. For this reason, making multiple PSF models over time may not be practical or useful.

Additionally, the PSF of the star changes with increasing magnitude. The brighter the star, the more pixels must be used to contain the flux. Even when the CCD does not saturate, brighter observations still take up more pixels and will have a larger, more elongated PSF than dimmer stars. It is worth examining if there are significant changes in the PSF model between stars on the brighter end of the spectrum versus the dimmer end. It is possible that changes in magnitude may affect the shape of the PSF more than the channel or location. However, it is still worth noting that there are not many uncontaminated files in some campaigns, so the range of magnitudes for which separate PSF models are constructed cannot be too narrow.

Interpolation of observations to the same scale is also imperative to refine the PSF model. Although the centers were able to be aligned at the sub pixel level due to increasing pixels per data scale, the lack of exact scaling between observations guarantees that this PSF is not exact. There may be some slight shifting, not only of the centers, but also of the boundaries of the PSFs themselves for each star. Stars with realistically larger PSFs may be scaled too small and vice versa.

Furthermore, the center locations may not be exact. The locations were matched with Gaia values, but these may not be the exact true center of the star. It may be

worth determining the center of the star through max flux, as was done for the models, and seeing how this aligns with the *Gaia* locations.

Significant work remains in refining the PSF models for *K2* before they can be properly applied to look for exoplanets in the data.

6.2 Applying PSF Models Bright Stars

The PSF models developed in this paper can be used to refine the light curves for bright stars processed with `halophot`. Due to the complexity of developing the PSF models, the application of these models to the subtraction of background stars from bright star observations fell outside the scope of this paper. The PSF models are still rudimentary and require more refinement, along with more research on PSF models in the different campaigns of *K2* before they are functional enough to use for subtraction of background stars around bright targets.

6.3 Using Background Stars

In addition to looking for exoplanets in the light curves from bright stars, the light curves of the extracted background stars could also be searched for exoplanets. These background stars are numerous. Each of these subtracted stars is a potential light curve with potential transits in the signal. Though these background stars are less flux-rich, the BLS method described above has been used to detect planets around dim stars.

Searching these extracted lightcurves opens the door to a multitude of new, never before observed stars. None of these background stars were intentional targets, and thus were not specific targets of interest.

Additionally, it is possible some of them may align with TESS targets, since the scope of *TESS* is much more complete. Matching these background targets to observations from *TESS* could provide a more complete flux timeseries for *TESS* targets. This

extra data may be able to help find exoplanet transits that were previously unobserved. Transits require repetition to be found and confirmed. A single transit could be a false positive, but with more transits of the same period, it becomes more and more likely that there is a planet at that period. These accidentally imaged background stars could provide the necessary data to help verify transits in *TESS* data as the mission continues. Possible transits found in these background stars could also be verified with *TESS* as the mission continues hopefully long past its intended timeframe, and become future *TESS* targets of interest.

6.4 Bright Stars in TESS

The current NASA exoplanet mission, *TESS*, while having a higher saturation threshold, does still have some stars that saturate. *TESS* is still in its first round of observations, viewing the sky in 26 sectors, 13 in the northern hemisphere and 13 in the southern hemisphere. *Kepler* was only able to observe a small portion of the sky, first trained on the single field of view near Cygnus and then only able to view along the path of the ecliptic as *K2*. *TESS*, on the other hand, will observe the entire sky. Many new exoplanets have already been found using the incoming data. With the inclusion of bright stars that saturate, there could be even more!

Bibliography

William J. Borucki, David Koch, Gibor Basri, Natalie Batalha, Timothy Brown, Douglas Caldwell, John Caldwell, Jørgen Christensen-Dalsgaard, William D. Cochran, Edna DeVore, Edward W. Dunham, Andrea K. Dupree, Thomas N. Gautier, John C. Geary, Ronald Gilliland, Alan Gould, Steve B. Howell, Jon M. Jenkins, Yoji Kondo, David W. Latham, Geoffrey W. Marcy, Søren Meibom, Hans Kjeldsen, Jack J. Lissauer, David G. Monet, David Morrison, Dimitar Sasselov, Jill Tarter, Alan Boss, Don Brownlee, Toby Owen, Derek Buzasi, David Charbonneau, Laurance Doyle, Jonathan Fortney, Eric B. Ford, Matthew J. Holman, Sara Seager, Jason H. Steffen, William F. Welsh, Jason Rowe, Howard Anderson, Lars Buchhave, David Ciardi, Lucianne Walkowicz, William Sherry, Elliott Horch, Howard Isaacson, Mark E. Everett, Debra Fischer, Guillermo Torres, John Asher Johnson, Michael Endl, Phillip MacQueen, Stephen T. Bryson, Jessie Dotson, Michael Haas, Jeffrey Kolodziejczak, Jeffrey Van Cleve, Hema Chandrasekaran, Joseph D. Twicken, Elisa V. Quintana, Bruce D. Clarke, Christopher Allen, Jie Li, Haley Wu, Peter Tenenbaum, Ekaterina Verner, Frederick Bruhweiler, Jason Barnes, and Andrej Prsa. Kepler planet-detection mission: Introduction and first results. *Science*, 327(5968):977–980, 2010. ISSN 0036-8075. doi: 10.1126/science.1185402. URL <https://science.sciencemag.org/content/327/5968/977>.

Stephen T. Bryson, Peter Tenenbaum, Jon M. Jenkins, Hema Chandrasekaran, Todd Klaus, Douglas A. Caldwell, Ronald L. Gilliland, Michael R. Haas, Jessie L. Dotson, David G. Koch, and et al. The kepler pixel response function. *The Astrophysical Journal*, 713(2):L97L102, Mar 2010. ISSN 2041-8213. doi: 10.1088/2041-8205/713/2/197. URL <http://dx.doi.org/10.1088/2041-8205/713/2/L97>.

Jeffrey Cleve, Steve Howell, Jeffrey Smith, Bruce Clarke, Susan Thompson, Stephen Bryson, Mikkel Lund, Rasmus Handberg, and William Chaplin. That’s how we roll:

- The nasa k2 mission science products and their performance metrics. *Publications of the Astronomical Society of the Pacific*, 128, 12 2015. doi: 10.1088/1538-3873/128/965/075002.
- Knicole Colon. K2 campaign 7 data available, Apr 2016. URL <https://keplerscience.arc.nasa.gov/k2-campaign-7-data-available.html>.
- Hans J. Deeg and Roi Alonso. Transit photometry as an exoplanet discovery method. *Handbook of Exoplanets*, page 633657, 2018. doi: 10.1007/978-3-319-55333-7_117. URL http://dx.doi.org/10.1007/978-3-319-55333-7_117.
- Brian Dunbar. Kepler’s field of view in targeted star field, 2008. URL https://www.nasa.gov/mission_pages/kepler/multimedia/images/fov-kepler-drawing.html.
- Martin Ester, Hans-Peter Kriegel, Jrg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. *AAAI*, page 226231, 1996. URL <https://www.aaai.org/Papers/KDD/1996/KDD96-037.pdf>.
- Michael Henderson. *High Precision Photometry of Faint White Dwarf Stars from K2 Data*. PhD thesis, Wesleyan University, 2019.
- Steve B. Howell, Charlie Sobeck, Michael Haas, Martin Still, Thomas Barclay, Fergal Mullally, John Troeltzsch, Suzanne Aigrain, Stephen T. Bryson, Doug Caldwell, William J. Chaplin, William D. Cochran, Daniel Huber, Geoffrey W. Marcy, Andrea Miglio, Joan R. Najita, Marcie Smith, J. D. Twicken, and Jonathan J. Fortney. The K2 Mission: Characterization and Early Results. *PASP*, 126(938):398, April 2014. doi: 10.1086/676406.
- John M. Jenkins. Kepler data processing handbook. Technical Report KSCI-19081-002,

- NASA, NASA Ames Research Center; Moffett Field, CA, United States, January 2017.
- Michele Johnson. Mission manager update: K2 in campaign 5, May 2015. URL <https://www.nasa.gov/ames/kepler/mission-manager-update-k2-in-campaign-5>.
- G. Kovács, S. Zucker, and T. Mazeh. A box-fitting algorithm in the search for periodic transits. *A&A*, 391:369–377, August 2002. doi: 10.1051/0004-6361:20020802.
- M. Libralato, L. R. Bedin, D. Nardiello, and G. Piotto. A psf-based approach to kepler/k2 data i. variability within the k2 campaign 0 star clusters m 35 and ngc 2158. *Monthly Notices of the Royal Astronomical Society*, 456(2):11371162, Dec 2015. ISSN 1365-2966. doi: 10.1093/mnras/stv2628. URL <http://dx.doi.org/10.1093/mnras/stv2628>.
- Mikkel N. Lund, Rasmus Handberg, Guy R. Davies, William J. Chaplin, and Caitlin D. Jones. K2p2a photometry pipeline for the k2 mission. *The Astrophysical Journal*, 806(1):30, May 2015. doi: 10.1088/0004-637x/806/1/30.
- Eli Maor and Ian Stewart. *To infinity and beyond: a cultural history of the infinite*. Princeton University Press, 2017.
- Ismael Mireles. *Searching for Exoplanets Around the Brightest Stars in K2*. PhD thesis, Wesleyan University, 2019.
- Michael Perryman. The History of Exoplanet Detection. *Astrobiology*, 12(10):928–939, Oct 2012. doi: 10.1089/ast.2011.0784.
- Benjamin J. S. Pope, Timothy R. White, Will M. Farr, Jie Yu, Michael Greklek-McKeon, Daniel Huber, Conny Aerts, Suzanne Aigrain, Timothy R. Bedding,

Tabetha Boyajian, Orlagh L. Creevey, and David W. Hogg. The k2 bright star survey i: Methodology and data release, 2019.

George R. Ricker, Joshua N. Winn, Roland Vanderspek, David W. Latham, Gspr . Bakos, Jacob L. Bean, Zachory K. Berta-Thompson, Timothy M. Brown, Lars Buchhave, Nathaniel R. Butler, R. Paul Butler, William J. Chaplin, David B. Charbonneau, Jrgen Christensen-Dalsgaard, Mark Clampin, Drake Deming, John P. Doty, Nathan De Lee, Courtney Dressing, Edward W. Dunham, Michael Endl, Francois Fressin, Jian Ge, Thomas Henning, Matthew J. Holman, Andrew W. Howard, Shigeru Ida, Jon M. Jenkins, Garrett Jernigan, John Asher Johnson, Lisa Kaltenegger, Nobuyuki Kawai, Hans Kjeldsen, Gregory Laughlin, Alan M. Levine, Douglas Lin, Jack J. Lissauer, Phillip MacQueen, Geoffrey Marcy, Peter R. McCullough, Timothy D. Morton, Norio Narita, Martin Paegert, Enric Palle, Francesco Pepe, Joshua Pepper, Andreas Quirrenbach, Stephen A. Rinehart, Dimitar Sasselov, Bunei Sato, Sara Seager, Alessandro Sozzetti, Keivan G. Stassun, Peter Sullivan, Andrew Szentgyorgyi, Guillermo Torres, Stephane Udry, and Joel Villaseñor. Transiting Exoplanet Survey Satellite. *Journal of Astronomical Telescopes, Instruments, and Systems*, 1(1):1 – 10, 2014. doi: 10.1117/1.JATIS.1.1.014003. URL <https://doi.org/10.1117/1.JATIS.1.1.014003>.

A. Santerne, R. F. Daz, J. M. Almenara, A. Lethuillier, M. Deleuil, and C. Moutou. Astrophysical false positives in exoplanet transit surveys: why do we need bright stars ?, 2013.

S. Seager and G. MallenOrnelas. A unique solution of planet and star parameters from an extrasolar planet transit light curve. *The Astrophysical Journal*, 585(2):10381055, Mar 2003. ISSN 1538-4357. doi: 10.1086/346105. URL <http://dx.doi.org/10.1086/346105>.

- O. Struve. Proposal for a project of high-precision stellar radial velocity work. *The Observatory*, 72:199–200, Oct 1952.
- Jeffrey Van Cleve and Doug Caldwell. Kepler: A search for terrestrial planets. kepler instrument handbook. Technical Report KSCI-19033-001, NASA, NASA Ames Research Center; Moffett Field, CA, United States, July 2009.
- Jeffrey Van Cleve and Kenneth Mighell. Kepler: A search for terrestrial planets. k2 handbook. Technical Report KSCI-19116-001, NASA, NASA Ames Research Center; Moffett Field, CA, United States, July 2019.
- T. R. White, B. J. S. Pope, V. Antoci, P. I. Ppacs, C. Aerts, D. R. Gies, K. Gordon, D. Huber, G. H. Schaefer, S. Aigrain, and et al. Beyond the kepler/k2 bright limit: variability in the seven brightest members of the pleiades. *Monthly Notices of the Royal Astronomical Society*, 471(3):28822901, Aug 2017. ISSN 1365-2966. doi: 10.1093/mnras/stx1050. URL <http://dx.doi.org/10.1093/mnras/stx1050>.
- Bo Xin, eljko Ivezi, Robert H. Lupton, John R. Peterson, Peter Yoachim, R. Lynne Jones, Charles F. Claver, and George Angeli. A study of the point-spread function in sdss images. *The Astronomical Journal*, 156(5):222, Oct 2018. ISSN 1538-3881. doi: 10.3847/1538-3881/aae316. URL <http://dx.doi.org/10.3847/1538-3881/aae316>.