Kepler Transits

Ethan Rooney

15 Dec 2018

Abstract

Using a python Library developed for working with Kepler data, I did a survey of Light curves in the Kepler Objects of Interest Data base. Specifically I was looking for objects with a planetary transit causing a dip in luminosity of the star system of more than 0.5%. As the luminosity decrease during a transit is proportional to the ratio of the host star and the exoplanets radius, to meet this threshold the exoplanet must have a radius at least one tenth the size off the host star. For reference, an occlusion of the sun by Jupiter as sufficiently large distances would account for ~1% dip in luminosity. Furthermore, due to the nature of Kepler quarterly survey, most of the objects analyzed have a relatively short periodicity. Practically this means that many of the exoplanets I have looked at have fallen into the "Hot Jupiter" category. With a few interesting anomalies as well.

Kepler

Kepler was a space based telescope in a sun centric orbit. It launched on March 7, 2009 (Metcalfe 2008) and was performed its intended function until the second failure of its reaction wheel in May of 2013 (Abdul-Masih et al. 2016). Kepler maintained a fix on a patch of sky ~105 degrees square, using 3 reaction wheels. When the first reaction wheel failed, the mission continued on 2 with, but with the failure of the second reaction wheel the original mission plan could not be continued, and the satellite's mission was rescoped and continued under the name K2.

Data Acquisition and Analysis

With 95 million pixels onboard to collect all of the information available would have overwhelmed the onboard storage and communication. Thus once Kepler was launched, NASA selected 5.4 million of the pixels to write to memory. These pixels where associate with ~200,000 of the stares located in the viewing window where selected for observation during which time the light from those stars was collected and a "postage stamp" of the light flux from that star was collected. This was achieved by taking an exposure every 6.5 seconds and summing the exposures over 29.4 minutes, this being referred to as long cadence. A subset of these stars of particular interest were selected for a "short cadence observation" these short cadence observations where used for objects of high interest and were usually observed over the course of a 30 day cycle before rotating out of the short cadence observation group. ("Kepler and K2 Data Products," n.d.)

Postage Stamps

In the image below, we have a typical Target Pixel file. This image is a long cadence observation of Kepler Object 008153411. We can see the primary object to be observed in the (3,2) sector of the Figure 1–index starting at (0 column, 0 row) i.e. the top left corner–and we have additional light spilling over into the surrounding pixels. At (1,0) there appears to be contamination coming from another star near by. In an instance like this we need to apply a mask to the data to minimise the noise in the image while losing as little of the transit signal as possible.

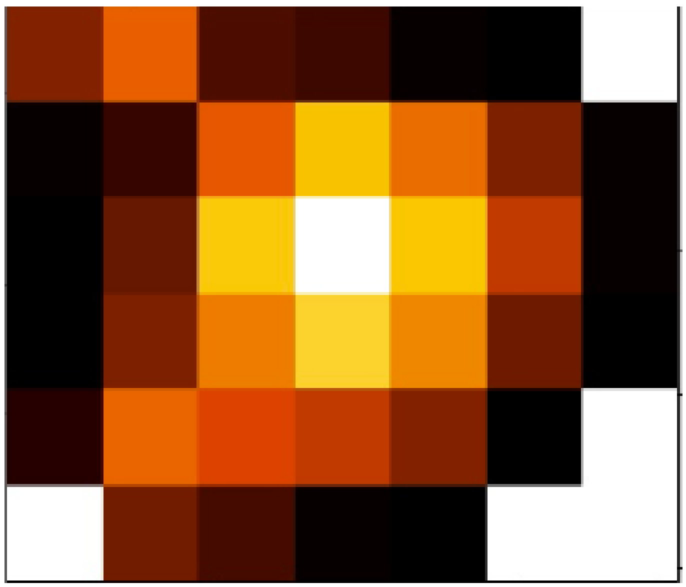

Figure 1: Sample Target Pixel File

Thus a mask as seen in Figure 2 is applied. The black pixels represent those no data for this object were gathered, the yellow field is all of the information for this TPF that are to be rejected, and finally the white represents the pixels which should be summed to calculate the raw light flux of the star.

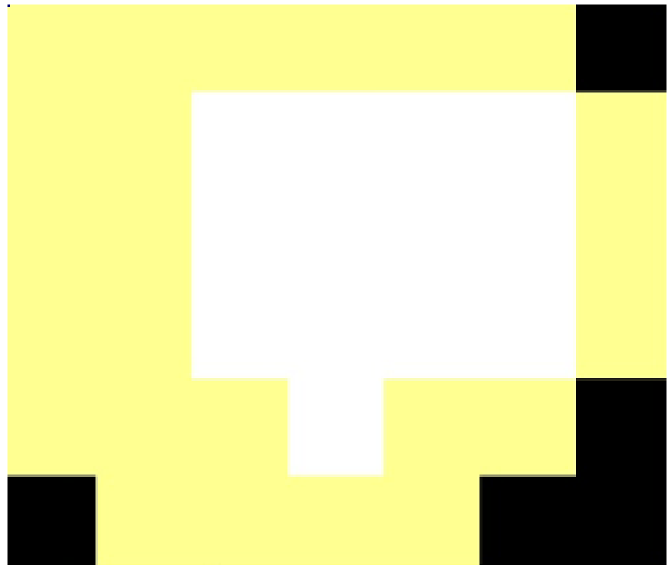

Figure 2: Masking the Target Pixel File

Once a mask has been formed for a Target Pixel File, (TPF), the unmasked pixels value can now be calculated, the raw value output is electrons, as the photon capture excites an electron across a diode. Each electron is counted to yield a raw intensity. This numerical value as well as the metadata such as time and location for each masked TPF is then concatenated into a data array, referred to as a light curve. This mask application, and concatenation mask selection, was performed using a library created specially to analyze Kepler data (Vinícius et al. 2018). The code below is an example of how to generate a light curve using lightkurve

```
tpf = KeplerTargetPixelFile.from_archive(star, //
                                          quarter, //
                                          quality_bitmask='hardest');
lc = tpf.to_lightcurve(aperture_mask='all');
```

The code snippet above does a few things, first it creates a variable **tpf** then it assigns that variable the value of a specified Kepler TPF located in the MAST database. With the values of star, quarter, and a bitmask of choice. 'star' and 'quarter' are themselves values of your choice. They can even be an array that is looped through so as to automate the production of light curves to analyse.

After a TPF for a given star and quarter are selected. The file is then passed into another function to_lightcurve(); this function then generates a light curve file that can then be analysed, like the one in Figure 3. As you can see from its y axis, I am no longer using a raw intensity score, rather I have plotted the light curve with its relative intensity, where the value is not in electrons, but rather a relative reverence with the average light output defined as 1.

While this light curve is of reasonable quality, we can still see major trends in the light curve unrelated to the transit of the exoplanet, not only that but the overall drift in the stars light curve would be enough obsucure the transit dip from any analyse we could run on the light curve in N time. To unobsure this transit we must remove any change in the value due to the regular change in light intensity of the star to get Figure 4 below. This flattening is performed by fitting a polynomial function to the prior graph and then finding the delta between that function and the light curve, then normaling the funtion back to an average of on.

From here we can see a regular structure appearing ~20 days. To check this we can perform a

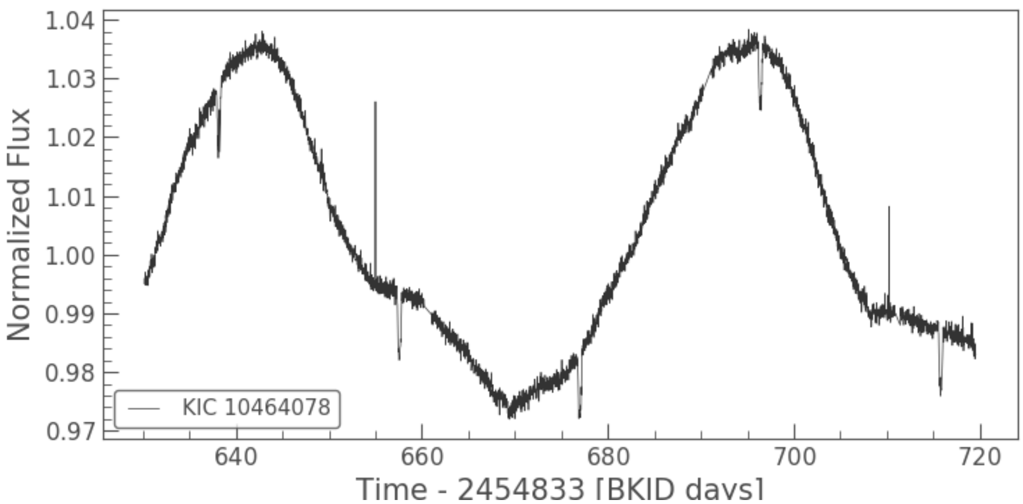

Figure 3: Light curve of Kepler Object 10464078 Quarter 7

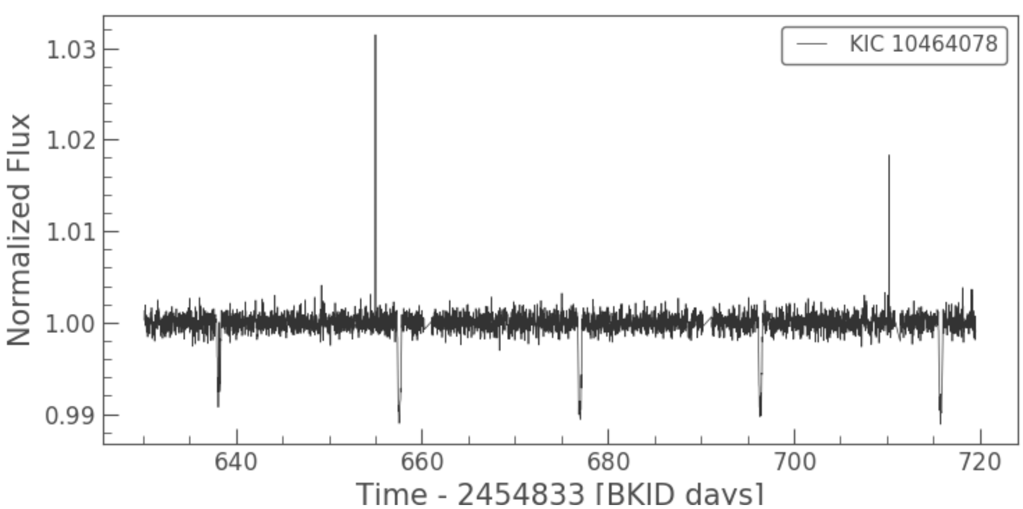

Figure 4: Flattened Light Curve of Kepler Object 10464078 Quarter 7

Lomb-Scargle transform on Figure 4. Resulting Figure 5.

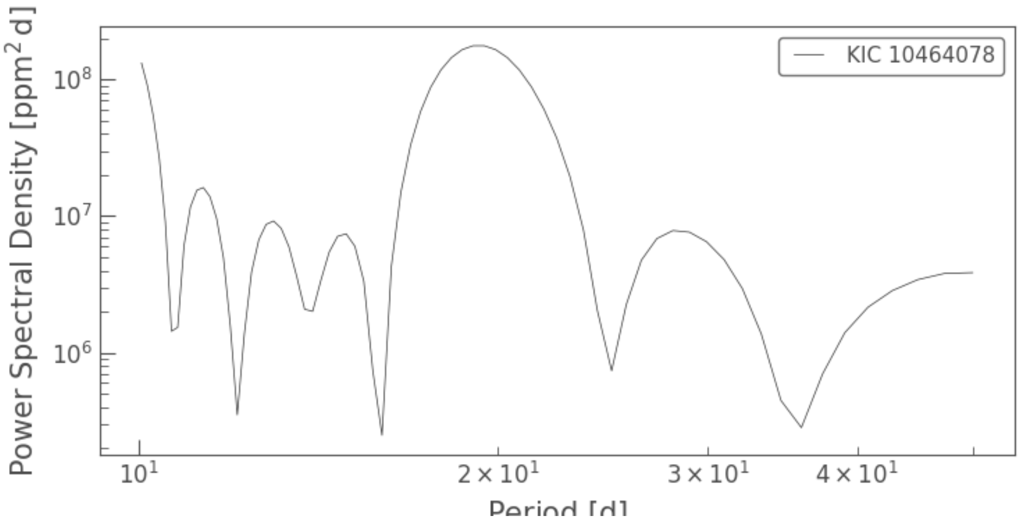

Figure 5: Lomb-Scarge Transform for Light Curve of Kepler Object 10464078 Quarter 7

Again in the periodogram we see a strong structure at approximately 20 days to conferm this we can find the highest power between 10 and 50 days.

```
p = lc_flat.to_periodogram(min_period=10*u.day, max_period=50*u.day, oversample_factor=10)
p.plot(format='period', scale='log')
plt.savefig('periodogram/{0}_quarter_{1}'.format(str(star),i))
period = p.period_at_max_power
print('Best period: {}'.format(period))
```

Which returned

```
19.4821015284
```

This returns the period with the strongest presence in the data, and we find a period. When we then fold the flattened curve over that period we with the code below. (Note: The work of VanderPlas -(VanderPlas 2018) has been invaluble in working this out, and his work needs to be recognized for its contribution)

```
lc_flat.fold(period.value).scatter();
plt.savefig('fold/{0}_quarter_{1}'.format(str(star), i))
```

We get Figure 6.

Extractable Physics

Now That we have gone through extracing a periodicity from a Kepler TPF, we can discuss what else can be gleaned from the given data. For KOI 10464078, we can lookup in the Kepler data base a star mass of 0.680 soler masses, and a radius of 0.657 soler radii. Using this we can begin to calculate the relative size of the planet.

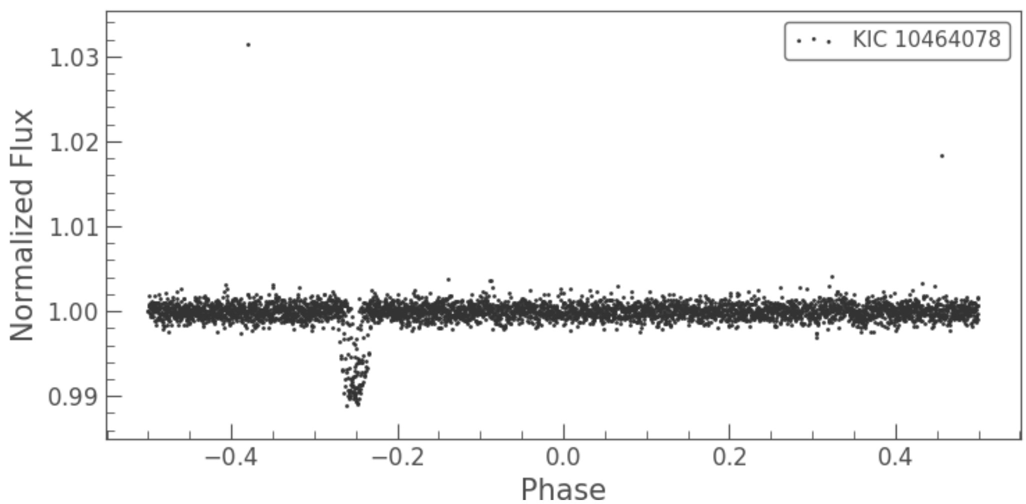

Figure 6: Folded Plot on a Period of ~ 19 Days for Light Curve of Kepler Object 10464078 Quarter 7

Exoplanet radius

First lets find the ratio of the planet to the star.

Let I be the normalized intensity of the stars light. $\implies I = 1$, and I_t be the intesity of the star durring a transit. And we will assume that both the star and the planet can be aproximated as uniformly lit disks.

Then $I \propto A_{star}$ and $I_t \propto A_{star} - A_{exoplanet}$ Where A_{star} is the area of the star disk and $A_{exoplanet}$ is the area of the exoplanet disk.

$\implies \frac{I_t}{I} = \frac{2\pi({r_{star}}^2 - {r_{exoplanet}}^2 i)}{2\pi r_{star}}$

$\implies I_t - 1 = -\frac{r_{exoplanet}}{r_{star}}$

$\implies \sqrt{\Delta I r_{star}^2} = r_{exoplanet}$

For KOI 10464078 we get a radius of 0.068 solar radii, or 47.3 Thousand km.("Kepler_KOI," n.d.)

Orbital Radius

To find the Orbital radius of the planet, we can use the Keplers satallite law.

$T = 2\pi\sqrt{\frac{r_{orbit}^3}{Gm_{star}}}$

Because we have all of the information other than the radius of the orbit we rearange to find:

$r_{orbit} = (\frac{T^2}{4\pi^2} Gm_{star})^{\frac{1}{3}}$

$\implies r_{orbit} = 9.452 * 10^6$ km

If we compare these values to those in the Kepler Mast Dataset NASA calculated the period to be 19.4 Days, with a ratio of 0.097 for the exoplanet radius to host star radius("Kepler_KOI," n.d.). Values quite close to those calculated by myself. Interesting to note that NASA has labeled this a False

Positive for an exoplanet. I have been unable to find out why they have declared this a false positive, as the signal was persistent and strong through all of the quarters for which Kepler data has been taken for this star.

Anomolies

Another interesting object I came across is Object 846285 in quarter 8 an unusual transit occured see in in Figure 7.

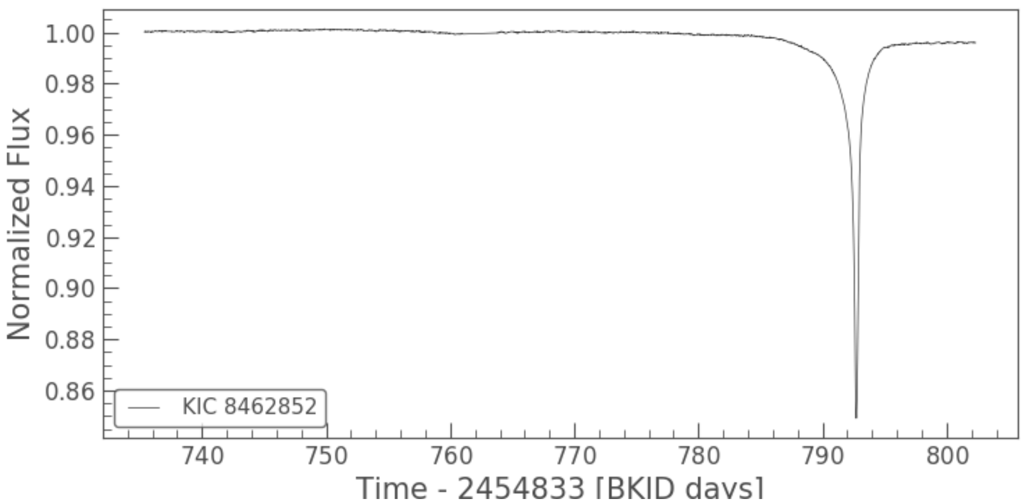

Figure 7: Anomolus Transit in Quarter 8

Here we have a massive change in a stars luminosity. But further observations of the object have not shown it to have a corrispondingly large mass. As the star does not seem to be pulled around a common center of gravity with the orbiting object. and though it has lead to lots of wild speculation as to the nature of the object. The actuall knowables are that it has a period of approximately 750 days and is probably made of a diffuse gas or loose rubble, possibly similer to what is contained in the Oort Cloud, but probably not Aliens(Tamblyn, n.d.).

Hot Jumping Jupiter

Continouing with the list of unsuall findings, we have Hot Jupitors. These are planets with ~0.25 Jovian Masses, and an orbrital period of ~10 days. As an example of that we have Figure 8

Here we can see a massive planet with a radius $\sim\frac{1}{5}$ The size of the host star with an orbital period $\frac{1}{40}$th the period of Mercury. This has challenged all prior assumptions about solar system formation.

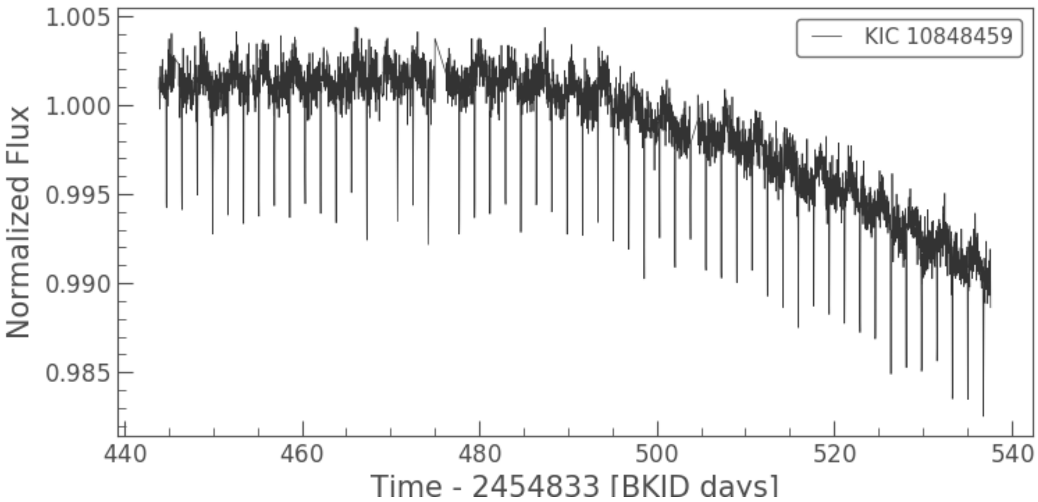

Figure 8: Hot Jupiter Object 10848459 Quarter 5

Source code for analyse

```
#!/bin/bash/python

from lightkurve import log, KeplerTargetPixelFile, Periodogram
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
import astropy.units as u
log.setLevel('WARNING')  # Ignore download messages

#This pandas Data frame contains the objects of interest to analize in a cvs
df = pd.read_csv('koi.csv', header=0)
#The analisis can be run on any number of objects in the Relevent table here I select what value to c
n=100
for a in range(n):
    # kepid is the header of the relevent column for extracting a Kepler ID
    star= df['kepid'][a]
    # Loop over all of the quarters of data for a given Kepler ID
    for i in range(17):
        #This handles any server timeout errors with out quitting, additionally if the object doesn't
        try:
            # This function will be removed from lightkuve soon.
            # This function takes input for a kepler objest and quarter, applies a bitmask to it and
            tpf = KeplerTargetPixelFile.from_archive(star, quarter=i, quality_bitmask='hardest');
        #Permit keyboard interupt from being gobbled up by the try error handling
        except KeyboardInterrupt:
            raise
        # on any error other than Keyboard interrupt, informs if Data requested is unavailable.
        except:
            print('no data for {0} quarter {1}'.format(str(star), i))
```

```
#If no error getting TPF Begins analysis of the TPF
else:
    print('analysing star {0} quarter {1}'.format(str(star),i))
    #converts the postage stamp of kepler data to a lightcurve
    lc = tpf.to_lightcurve(aperture_mask='all');
    #Applies a default flattening function to the light curve to removce systemic changes lik
    lc_flat = lc.flatten()
    #Allows for easy referance to the time and flux
    time, flux = lc_flat.time, lc_flat.flux
    #find the diff between min and max flux and if it is big enough do something with it.
    delta_flat = ((np.mean(flux)-min(flux))/np.mean(flux))
    print(delta_flat)
    # If the diffence between the max and min in the flattened curve is more than 0.5% we wan
    if delta_flat > 0.005:
        #Preserves the original light curve for further observation
        lc.plot()
        #outputs to file with the name and quarter of the object
        plt.savefig('raw/{0}_quarter_{1}_uncorrected.png'.format(str(star), i))
        #Same thing but for the flattened curve
        lc_flat.plot(linestyle='solid');
        plt.savefig('flat/{0}_quarter_{1}_flat.png'.format(str(star), i))
        #Outputs a periodogram for observation
        p = lc_flat.to_periodogram(min_period=5*u.day, max_period=50*u.day, oversample_factor
        p.plot(format='period', scale='log')
        plt.savefig('periodogram/{0}_quarter_{1}'.format(str(star),i))
        period = p.period_at_max_power
        print('Best period: {}'.format(period))
        #Finally outputs a folded curve for observation.
        lc_flat.fold(period.value).scatter();
        plt.savefig('fold/{0}_quarter_{1}_period{3}'.format(str(star), i,period))
        #Clear all plots from memory, oterwise this will eat your memory
        plt.close('all')
```

A Special Thanks to Mike Moss and Nick Gorgone, with out there help, I would not have analysed nearly as much data.

P.S. I have ~100 GB of data on my machine Let me know If you want any of it or if it would help my grade.

Regards Ethan Rooney

References

Abdul-Masih, M., A. Prša, K. Conroy, S. Bloemen, T. Boyajian, L. R. Doyle, C. Johnston, et al. 2016. "Kepler Eclipsing Binary Stars. VIII. Identification of False Positive Eclipsing Binaries and Re-extraction of New Light Curves" 151 (April): 101. doi:10.3847/0004-6256/151/4/101.

"Kepler and K2 Data Products." n.d. https://keplerscience.arc.nasa.gov/data-products.html.

"Kepler_KOI." n.d. https://archive.stsci.edu/kepler/koi/search.php.

Metcalfe, Travis S. 2008. "M. Dikpati et Al. Asteroseismology and the Solar-Stellar Connection."

Tamblyn, Thomas. n.d. "Alien Megastructure." https://www.huffingtonpost.co.uk/2015/11/30/alien-megastructure-solar-system-spotted-by-kepler-space-telescope-most-likely-a-star-swarmed-by-comets-say-nasa_n_8679858.html.

VanderPlas, Jacob T. 2018. "Understanding the Lomb–Scargle Periodogram." *The Astrophysical Journal Supplement Series* 236 (1): 16. http://stacks.iop.org/0067-0049/236/i=1/a=16.

Vinícius, Zé, Geert Barentsen, Christina Hedges, Michael Gully-Santiago, and Ann Marie Cody. 2018. "KeplerGO/Lightkurve." doi:10.5281/zenodo.1181928.