Multiwavelength Photometry Derived from Monochromatic Kepler Data

Christina Hedges1,2, Rodrigo Luger3, Jessie Dotson2, Daniel Foreman-Mackey3, and Geert Barentsen1,2
1 Bay Area Environmental Research Institute, P.O. Box 25, Moffett Field, CA 94035, USA
2 NASA Ames Research Center, Moffett Field, CA, USA
3 Center for Computational Astrophysics, Flatiron Institute, New York, NY, USA

Received 2020 September 15; revised 2020 December 9; accepted 2020 December 10; published 2021 January 29

Abstract

The Kepler mission has provided a wealth of data, revealing new insights in time-domain astronomy. However, Kepler’s single bandpass has limited studies to a single wavelength. In this work we build a data-driven, pixel-level model for the pixel response function (PRF) of Kepler targets, modeling the image data from the spacecraft. Our model is sufficiently flexible to capture known detector effects, such as nonlinearity, intrapixel sensitivity variations, and focus change. In theory, the shape of the Kepler PRF should also be weakly wavelength-dependent, due to optical chromatic aberration and a wavelength-dependent detector response functions. We are able to identify these predicted changes in shape of the PRF using the residuals between Kepler data and our model. In this work, we show that these PRF changes correspond to wavelength variability in Kepler targets using a small sample of eclipsing binaries. Using our model, we demonstrate that pixel-level light curves of eclipsing binaries show variable eclipse depths, ellipsoidal modulation, and limb darkening. These changes at the pixel level are consistent with multiwavelength photometry. Our work suggests that each pixel in the Kepler data of a single target has a different effective wavelength, ranging from ≈550 to 750 nm. In this proof of concept, we demonstrate our model, and discuss possible uses for the wavelength-dependent PRF of Kepler. These uses include characterizing variable systems, and vetting exoplanet discoveries at the pixel level. The chromatic PRF of Kepler is due to weak wavelength dependence in the optical systems and detector of the telescope, and similar chromatic PRFs are expected in other similar telescopes, notably the NASA TESS telescope.

Unified Astronomy Thesaurus concepts: Time series analysis (1916); Multi-color photometry (1077); Astronomical instrumentation (799)

1. Introduction

The Kepler, K2, and TESS missions have demonstrated the power of time-series photometry. Each of these missions has performed its observations in a single wide bandpass in order to increase signal-to-noise ratio and reduce complexity of the instrument. This strategy has enabled exquisite time-series photometry. Data from these missions have been used to identify exoplanet candidates (e.g., Thompson et al. 2018), unlock the inner workings of red giant stars (e.g., Bedding et al. 2011), unravel the history of our Galaxy (e.g., Silva Aguirre et al. 2018), and elucidate the early behavior of supernovae (e.g., Dimitriadis et al. 2019), among many other accomplishments. Beyond single-wavelength observations, observing the color and time-series variability of astronomical objects simultaneously has the potential to unlock new insights about astronomical phenomena.

The NASA Kepler spacecraft has provided photometry of >400,000 targets over its 9 yr mission in a single bandpass. In this paper, we discuss how the Kepler data contain a weak wavelength dependence, due to the optical design and detector response. This weak wavelength dependence causes a “chromatic pixel response function,” which can be used to reveal multiwavelength photometry of astronomical targets in the Kepler catalog by extracting pixel-level light curves. Once systematic effects from the detector can be modeled, the weak effect of wavelength dependence in each pixel can be revealed in the residuals. This chromatic information can unlock varied new inferences using the wavelength variability of the Kepler sample.

In this paper we demonstrate and validate the wavelength dependence of the Kepler PRF. In Section 1 we introduce the characteristics of the Kepler telescope. In Section 2 we present a data-driven model that can be used to model the Kepler PRF, including wavelength-independent, common instrument systematics. Using this model we correct instrument systematics in each Kepler pixel independently, leaving any wavelength dependence intact. We use these pixel-level light curves to identify weak wavelength dependence in Kepler pixel data. In Section 3 we demonstrate that the wavelength dependence of the Kepler PRF is detectable by using literature examples of characterized, bright eclipsing binaries. We also present a coarse calibration of the effective wavelengths probed by the Kepler data. In Section 5 we discuss applications of this model and how it can be used for a wide range of astrophysical inferences and for vetting planet candidates at the pixel level. We discuss limitations and future improvements of the model.

1.1. Characteristics of the Kepler Spacecraft

During the prime mission (2009–2013), Kepler provided stable, precise time-series measurements of ≈200,000 targets, each with high-precision photometry lasting up to 4 yr. These data capture time-series variability at extreme precision; Kepler achieved precision better than 30 ppm for Kepler band magnitudes less than 12.5 over 6.5 hr (Gilliland et al. 2011). Kepler is capable of observing only at a single wavelength. The center of the Kepler bandpass is ≈640 nm, and it has 50% relative response between 450 and 825 nm. More information on the characteristics of the Kepler telescope can be found in the Kepler Instrument Handbook (Van Cleve & Caldwell 2016), hereafter referred to as KIH. Kepler observations are split into quarters, where each target is observed continuously for three months. The spacecraft rotates each season, such that the target...
falls on a new detector module each season (every fourth quarter the star falls on the same module). During the prime mission the pointing was stable, with pointing better than an arce second over the duration of a quarter. This small, long-term drift is caused by velocity aberration (see Jenkins et al. 2010a) and focus change due to changes in spacecraft temperature. These factors cause the position of a target to vary slowly over the course of an observation, characteristically moving \( \lesssim 1 \) pixel over the course of a single quarter. These drifts, coupled with the intrapixel sensitivity variations (see KIH), cause the flux in each pixel to vary slowly over time.

The data volume from Kepler was high, owing to its short observing cadence (\( \approx 30 \) minutes in long-cadence mode, \( \approx 1 \) minute in short cadence mode). In order to reduce the data volume, the Kepler telescope downlinked data in small image cutouts, centered around targets of interest. Data were delivered in target pixel files (TPFs), which consist of stacks of images typically \( \approx 5 \times 5 \) pixels across, with one image per 30 minutes cadence. An example of a frame from a TPF is shown in Figure 1. TPFs are then converted into light curves by the Kepler Pipeline (Jenkins et al. 2010b) using simple aperture photometry (SAP). Pixels inside apertures specified by the Kepler Pipeline are summed in order to create a single flux time series of each target.

When flux from a target goes through the optical system of the telescope, it is spread from a point source into a broad point-spread function (PSF). The PSF for Kepler is wider than a single pixel, and so light from the source is registered on multiple pixels (for an example of the theoretical PSF of Kepler see Figure 2). The image that is recorded on the detector is known as the pixel response function (PRF), which is the true image combined with the detector sensitivity. For the bright target in Figure 1 the PRF covers tens of pixels.

During the prime mission, systematics in the light curves produced by the mission (such as focus change and velocity aberration) were corrected by the Kepler pipeline (see Kepler Data Processing Handbook, Jenkins 2017). The pipeline approach was to use cotrending basis vectors (CBVs) to remove common trends in the light-curve sample (see Smith et al. 2012). In the mission’s second phase, known as K2 (2013–2018) (Howell et al. 2014), the observations were subject to increased motion noise, owing to the spacecraft pointing in a two-reaction wheel mode. Intrapixel sensitivity variations cause the K2 time series to have variability at a level of \( \approx 0.1\% \), and were not explicitly corrected by the Kepler pipeline. The CBVs employed by the pipeline removed common systematics between targets, but owing to the unique sensitivity variations of each pixel, this approach is not able to completely remove the systematic motion noise. Several groups produced techniques to mitigate the motion noise in the light curves produced by the K2 mission (e.g., see Vanderburg & Johnson 2014; Aigrain et al. 2015; Luger et al. 2016), employing techniques to correct the pixel data local to the source. These works are able to achieve light curves with a precision within \( \approx 1–4\times \) that of the prime mission. Since any wavelength dependence in Kepler is predicted to be weak, it is paramount to correct all other known systematics in the Kepler data, such as motion over the intrapixel flat field. In this work, we will use similar techniques to those used to correct K2 systematics in Luger et al. (2016) to model systematics from data from the Kepler prime mission at the pixel level, leaving any wavelength dependence intact.

1.1.1. Saturation

The Kepler detector was designed to ensure that all flux from targets was captured, even in the case where targets would saturate the detector. The saturation limit for the detector is \( \approx 1.5 \times 10^5 \ e^-/s \). The Kepler detector is designed such that when a pixel reaches this limit, the flux “bleeds” vertically in columns above and below that pixel, creating a “bleed
column.” This ensures that the total flux of the target can still be measured, even for targets brighter than the saturation limit. This alters the shape of the PRF measured by the detector. Saturation causes unpredictable effects in the pixel-level data (e.g., exacerbated systematics from moiré patterns and undershoot), which are discussed in detail in the KIH. Pixels that are at the ends of bleed columns are also highly variable, as flux spills from all the pixels inside the column. In this work, saturated pixels, and pixels at the ends of bleed columns, are removed and not modeled.

1.2. Known Color Dependence of the Kepler PRF

There are two known effects that cause a wavelength dependence in the PRF of the Kepler telescope: (1) chromatic aberration from the optical system; (2) wavelength dependence of the intrapixel sensitivity. Together, these two effects combine to make the PRF of the instrument weakly color-dependent. Below we discuss each of these effects.

1.2.1. PSF Wavelength Dependence

An example of the theoretical, pre-flight PSF is shown in Figure 2, which demonstrates the chromatic aberration. This figure is adapted from the KIH (Van Cleve & Caldwell 2016). The PSF is shown modeled at two different wavelengths at the edges of the Kepler bandpass. Figure 2 shows that, in theory, the PSF shape should vary between these wavelengths. The true PSF is integrated over the Kepler bandpass which spans 430–830 nm, and so any color variation between the PSFs of blue and red targets is, in reality, difficult to detect. When the PSF is measured (creating a PRF), the shape also varies based on factors such as telescope motion, intrapixel sensitivity, spacecraft temperature, etc. Given these factors, it is difficult to identify these small shape differences in the real data.

The PSF models shown in Figure 2 are for qualitative purposes. While we can expect some color dependence in the instrument PSF, we do not have an estimate of the magnitude of the effect based on these models. Pre-flight models of the chromatic effects of the Kepler PSF, such as those used to generate the original data behind Figure 2, are not publicly available. Such PSF models are likely to have limited use to analyze the color dependence of the Kepler data, because they are based on theoretical models and do not accurately take into account the detector sensitivities. Also, since these models were generated by the mission on a coarse grid of points around the focal plane, and at a limited set of focus offsets, they would not capture the effect of PSF shape change across a channel.

1.2.2. Wavelength Dependence of Intrapixel Sensitivity

Each pixel in the Kepler detector is not uniformly sensitive. Kepler pixels tend to be most sensitive toward the center and less sensitive toward the edges (for further details, see Bryson et al. 2010). This effect causes a significant systematic for both Kepler and K2 data. Vorobiev et al. (2018) measured the intrapixel sensitivity using a Kepler flight spare of the detector, and found that the there is significant variation in the pixel response function at 450, 600, and 800 nm. While the same general trend of high sensitivity toward the center of each pixel holds for all wavelengths, the drop-off in sensitivity at the edges of the pixel is much steeper at shorter wavelengths. A light-curve measurement of a target that moves on the detector (due to either spacecraft motion, focus change, or velocity aberration) will have a variable flux due to sensitivity variation. Owing only to the wavelength dependence of the sensitivity noted by Vorobiev et al. (2018), this variability will have a steeper gradient at shorter wavelengths than at longer wavelengths. Similarly, a target that is stationary but undergoes wavelength-dependent variability will also have a slightly different response depending on where in the pixel it fell. Even with a completely achromatic PSF, the measured PRF of the instrument will have a wavelength dependence based on this effect alone.

1.2.3. The PRF

The true wavelength dependence of the PRF is the combination of both of the effects discussed here. Based on our theoretical understanding of the Kepler PSF and physical characterization of the detector, we can expect a wavelength dependence in the pixel data from Kepler. However, it is not possible based on theory alone to understand the magnitude of these effects. In this work, we develop a model for the instrument systematics of motion and focus change in each pixel, and leave intact any shape variation due to wavelength-dependent variability. Once the wavelength-independent instrument systematics are corrected using our model, we use the residuals to search for shape variation that correlates with astrophysical variability, which we attribute to color dependence.

2. Data-driven Model for Pixel Systematics

As discussed in Section 1.2, there is no reliable theoretical model for the color dependence of the Kepler PRF. In this work we will instead build a data-driven model for the shape of the PRF, which is flexible enough to capture instrument systematics such as focus change and velocity aberration. We use similar techniques to those applied in Luger et al. (2016) and Vanderburg & Johnson (2014). Each pixel is modeled independently. We will use this data-driven model to correct the systematics for each pixel individually. We will then identify cadences where there are significant residuals in multiple pixels, indicating that there are remaining systematics that are unaccounted for. Where all pixels exhibit a strong deviation from our systematics model, we attribute it to a PRF shape change. Using targets where there is a well known and predictable color change (i.e., eclipsing binaries), we will then show that this PRF shape change is due to wavelength-dependent variability. In the following sections we discuss the model for pixel-level systematics.

2.1. General Model

In this section we describe the data-driven model we use to model each pixel light curve, first in a general sense and then in the specific case of Kepler data in Section 2.2.

In general, we have a time-series measurement of a source, which we label \( f \), which contains both astrophysical time series and unwanted instrument systematics. In this general section, this could be any flux time-series data; \( f \) could be a single-pixel time series or a SAP light curve. Each data point in the time series has some associated error \( \sigma_f \). \( K_f \) is the covariance matrix for the data.

We label the true, astrophysical flux time series of the source as \( s \). \( s \) is a vector with \( n \) time points, and our goal is to extract the most reliable estimate of \( s \). We have measurements of the source from a telescope, which recorded the source brightness.
For this data set, an acceptable model would describe the underlying true astrophysical signal $s$, and some multiplicative factors and/or additive factors that take into account systematic trends from the telescope. For example, systematic motion over the intrapixel sensitivity function would cause flux to be altered by a multiplicative factor. An unrelated background source would cause an additive offset in the measured flux. The model for our measured signal $f$ is given as

$$f_m = S \cdot X \cdot w$$  \hspace{1cm} (1)$$

where $f_m$ is the model flux, $S$ is a diagonal matrix where the diagonal values are $s$, $X$ is a design matrix with predictors for systematics, and $w$ is a vector of weights for these systematic components.

The design matrix $X$ is a matrix made up of $m$ column vectors, each of length $n$ (the number of time points in the time series $f$). Each vector in $X$ is some predictor of a systematic, such as spacecraft temperature, motion, measured background scattered light etc. These vectors encode our understanding of the systematics. An example of the structure of $X$ is

$$X = \begin{pmatrix} x_{1,1} & x_{1,1} & \ldots & x_{1,m-1} & 1 \\ x_{2,1} & x_{2,2} & \ldots & x_{2,m-1} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{n,1} & x_{n,2} & \ldots & x_{n,m-1} & 1 \end{pmatrix} \hspace{1cm} (2)$$

Note that in the example matrix, the final column of $X$ is a row of ones. This column allows us to fit an additive offset in the time series (i.e., to fit the mean level).

In order to calculate the best fitting model, if we know the underlying astrophysical signal $s$, we must solve for the weights $w$. In this general case, we assume we know or have a reasonable estimate of $s$. To solve for $w$ and find the best fitting model $m$ we can use linear regression, employing the following relations (under the assumption that our errors are Gaussian):

$$K_w^{-1} = (S \cdot X^\top) \cdot K_f^{-1} \cdot (S \cdot X) \hspace{1cm} (3)$$

$$\hat{w} = K_w^{-1} \cdot ((S \cdot X^\top) \cdot K_f^{-1} \cdot f) \hspace{1cm} (4)$$

where $\hat{w}$ is a vector of length $m$ of the mean of the best fitting coefficients and $K_w$ is the covariance matrix of $w$ with size $m \times m$. These relations are derived in Luger et al. (2017). By rearranging Equation (1) we can now build a systematics model and find our best estimate of $s$, the true astrophysical signal.

2.2. Applying the Model to Kepler

In the more realistic case, we have several flux time series for the source, because the flux has been recorded on several pixels due to the broad PRF. We describe the flux time series for the $i$th pixel at time $t$ as $f_{i,t}$. We will assume there are $n$ points in time, $m$ vectors in our design matrix, and $l$ pixel time series in our TPF. In this work, we assume there is no covariance either between data points in time or between pixels, and our matrix $K_f$ is a simple diagonal matrix, where the diagonal values are fixed at $\sigma_{f}$ (the error on each flux measurement). This assumption implies that there are no correlations between either pixels or time points in the data set, and is adopted here for simplicity. A more robust analysis might include a $K_f$ with off-diagonal terms to include covariance in the time series (encoding covariance between “frames”) and covariance between pixels (encoding prior knowledge of the “true” PRF shape). In this first proof of concept, we treat each flux measurement as independent.

In this case, our “best guess” for the true underlying signal $s$ is the SAP light curve. This light curve is the sum of all pixels within the aperture predetermined by the Kepler pipeline. Our assumption here is that each pixel should contain the same signal as the “total” brightness from the source. Said another way, the $s$ measured in each pixel should be identical (for now, ignoring wavelength dependences). The SAP flux of a target is simply

$$s = \sum_{i=0}^{i'=l} f_{i,t}$$  \hspace{1cm} (5)$$

where $\sum_{i=0}^{i'=l}$ denotes summation of the pixel time series inside the optimum aperture specified by the Kepler Pipeline. (Note that $l$ denotes the pixels inside the aperture, and there are a total of $l$ pixels in the TPF). If there is no PRF shape variability in our data, the SAP light curve should be identically reflected in every pixel. Because we know there are some unrelated shape and position changes due to focus change, intrapixel sensitivity variations, and velocity aberration, we allow for common systematics between pixels with variable weights using the design matrix $X$ and weights $w$.

We build the design matrix $X$ using the following components:

1. The first two cotrending basis vectors (CBVs) from the Kepler Pipeline. These are components built by the Kepler pipeline (Jenkins et al. 2010b) and reflect common variability across all targets in a single Kepler channel. In our case, they are mostly predictive of focus change. Focus change causes a shape change in the PRF that is common across the entire channel, and has a greater magnitude at the start of quarters (where the spacecraft is changing temperature due to a new pointing). We include these components here, and they are considered a nuisance parameter. If these components are omitted, the PRF shape change due to focus change would not be fit well in our model. We find empirically in our work that the first two CBVs are the most predictive of focus change and sufficiently capture this variability.

2. The pixel column and row positions of the PSF as estimated by the Kepler Pipeline. These are provided by the TPF in the POS_CORR1 and POS_CORR2 arguments respectively, found in the pipeline processed fits files. We include a 2D, second-order polynomial in vectors of row ($r$) and column ($c$) position of the target. These vectors are $r, c, r^2, c^2, rc, r^2c, rc^2, (rc)^2$. This allows us to take into account the motion over the subpixel flat field. For Kepler, the motion takes a simple form and changes gradually, and so a low-order 2D polynomial is a reasonable model. For the data taken during the K2 mission, this assumption does not hold, and a different model for centroid position is required (see Section 5.3.1). This polynomial is analogous to the “arc-length” used in Vanderburg & Johnson (2014) to model K2 systematics.

3. A higher-order polynomial of flux and position. As shown in other works on the sensitivity of the Kepler pixels, the drop-off in sensitivity toward the edge of the pixel is not linear (e.g., see Vorobiev et al. 2018). As such, we cannot model the systematics linearly with $s$,
An example of the design matrix is

\[
X = \begin{bmatrix}
CBV_{0,0}, & CBV_{0,1}, & r_0, & c_0, & \cdots & 1 \\
CBV_{1,0}, & CBV_{1,1}, & r_1, & c_1, & \cdots & 1 \\
CBV_{2,0}, & CBV_{2,1}, & r_2, & c_2, & \cdots & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
CBV_{N,0}, & CBV_{N,1}, & r_N, & c_N, & \cdots & 1
\end{bmatrix}
\]

By utilizing Equations (3) and (4) we can now find the weights in each pixel \( w \) by using the relations

\[
K_w^{-1} = (S \cdot X)^\top \cdot K_f^{-1} \cdot (S \cdot X)
\]

(7)

and

\[
\hat{w}_i = K_w^{-1} \cdot ((S \cdot X)^\top \cdot K_f^{-1} \cdot f_i)
\]

(8)

where \( \hat{w}_i \) are the best fitting weights in pixel \( i \), and is a vector with length \( m \), with one coefficient per weight.

Our full model \( f_{\text{all}} \) (given by Equation (1)) is a matrix of size \( n \times l \), where the model has been computed for each time \( t \) in each independent pixel \( i \). The resultant model is a stack of “images” of our data-driven model, including instrument systematics, under the assumption that the astrophysical signal \( s \) is the same in each pixel, i.e., a model of the PRF, where the PRF shape may vary only in a way correlated with the CBV components or with the two-day basis-spline components. An example of the design matrix for this component are \( \{ s, r, c, s, r, c \} \).

4. A two-day B-spline. In order to capture time variability on short and long timescales, such as from the known Kepler rolling band (short term) and velocity aberration (long term), we include a third-order basis spline in our matrix that has knots every two days, this allows us to flexibly capture variability. In this work, where we are concerned with variability due to eclipsing binaries that occurs on timescales of hours, a two-day B-spline is acceptable. However, in other applications, a longer baseline may be more appropriate.

5. A column of ones. This allows the mean of the time series to be fit by the model. This column allows each pixel to have a different mean. In practice, this enables an offset in each pixel due to background flux from neighboring contaminant stars. This term takes into account the missing offsets from the other polynomial terms given above.

An example of the design matrix is

\[
X = \begin{bmatrix}
CBV_{0,0}, & CBV_{0,1}, & r_0, & c_0, & \cdots & 1 \\
CBV_{1,0}, & CBV_{1,1}, & r_1, & c_1, & \cdots & 1 \\
CBV_{2,0}, & CBV_{2,1}, & r_2, & c_2, & \cdots & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
CBV_{N,0}, & CBV_{N,1}, & r_N, & c_N, & \cdots & 1
\end{bmatrix}
\]

Finding \( w_i \) by jointly fitting multiple pixels and multiple targets, taking into account covariances between pixels, would enable us to find a model for the full detector, creating a data-driven PRF model in absolute terms. Such a model, given an input flux value for a given pixel and time \( f_{i,t} \), would be able to return the estimated flux for a given pixel based on the full context of the detector, accounting for all common systematics. However, to create such a model would require the incorporation of further information into the model, including the true underlying image (i.e., a sky catalog of the star field, including exact, subpixel positions of every target). Such a model is not necessarily easily described by simple linear combinations of polynomials and would be computationally intensive to generate. Generating and testing such a model is beyond the scope of this initial work, but would in theory enable this model to be expanded to the full Kepler sample.

2.4. Model Example: Kepler-1b

To demonstrate the detrending power of the model we have described, we apply it to the data set on transiting exoplanet Kepler-1b (also known as TrES-2b) (O’Donovan et al. 2006). Kepler-1b is an ideal test case for this model; the deep and short-period transits, as well as the bright host star, make the astrophysical signal in each pixel strong. In addition, a transiting planet is expected to cause no wavelength-dependent variability in the target flux. Even in the case of the hot planet of Kepler-1b, any emission from the planet would peak well outside the Kepler bandpass in the infrared, and changes in limb darkening are anticipated to be small. Figure 4 shows the model fit to a single pixel of Quarter 13 data of Kepler-1b. Our model has an exceptional fit to the data, providing an accurate fit at better than 200 ppm in most cadences. Our model is similarly accurate in each pixel in the data set.

The upper left panel of Figure 5 shows the SAP flux for Kepler-1b during Quarter 13 of Kepler, folded at the orbital period of Kepler-1b. Our pixel-level model is also summed over every pixel to create a model SAP flux light curve (blue). In the lower panel, the reduced chi-squared statistic is shown for each cadence, showing that our model has good agreement at all phases of the light curve (reduced chi-squared value of...
There are no significant excesses or deviations from the model during transit, and crucially there is no correlation between the residuals and the transit signal. A single quarter has been chosen for demonstration, but our model is similarly effective in all quarters.

2.5. Model Example: KIC 2708156

The Kepler-1b data set is not expected to have any wavelength dependence, since exoplanets are not bright enough to have significant emission in the Kepler bandpass. To find targets where wavelength dependence is both expected and predictable, we can look to eclipsing binaries. Eclipsing binaries are highly variable, and their astrophysical signal can easily be identified using only a single pixel of the broad PRF.

Eclipsing binaries also have a predictable color change; if the two stars have different temperatures, they will have different apparent colors in the Kepler bandpass. As one star eclipses the other, the color of the target will change, as each star’s contribution is downweighted when it is eclipsed. KIC 2708156 is an eclipsing binary, which was observed by Kepler during the prime mission (parameters available in Table 1). The right panel of Figure 5 shows our model fit to Quarter 13 data on KIC 2708156, and shows that our model has a significantly poorer fit during the eclipses of this target. The true SAP flux is shown for KIC 2708156, as well as the modeled SAP flux. The reduced chi-squared of the model fit is shown in the lower panel, summed up over all pixels in the aperture. The reduced chi-squared shows poor agreement between the model and the data.
Table 1
Table of Parameters of the Eclipsing Binaries Used in This Work. The Kepler Magnitudes and Effective Temperatures have been Derived from Matson et al. (2017). Here Subscript of 1 Denotes the Primary Body, and Subscript of 2 Denotes the Secondary. The Radius, Inclination and Amplitude Values were Derived in this Work by Fitting an Eclipsing Binary Model Using starry (Luger et al. 2019).

| KIC ID    | Kepler Mag | $T_{\text{eff}1}^a$ | $T_{\text{eff}2}^a$ | Period days | $t_0$ | Radius$_1$ | Radius$_2$ | Inclination$^b$ | Amplitude$_1$ | Amplitude$_2$ |
|-----------|------------|---------------------|---------------------|-------------|------|-------------|-------------|----------------|---------------|---------------|
| 2708156   | 10.672     | 11061               | 5671                | 1.891269    |      | 2.27939     | 2.08991     | 82.61361       | 0.90881       | 0.09119       |
| 3241619   | 12.524     | 5715                | 4285                | 1.703347    |      | 1.06836     | 1.02367     | 85.17719       | 0.76921       | 0.23079       |
| 4574310   | 13.242     | 7153                | 4077                | 1.306220    |      | 1.49917     | 1.58066     | 78.49502       | 0.80368       | 0.19632       |
| 4665999   | 13.016     | 7559                | 6846                | 2.248068    |      | 1.84602     | 1.36237     | 81.94781       | 0.74771       | 0.25229       |
| 4848423   | 11.825     | 6239                | 6176                | 3.003646    |      | 1.50426     | 1.38874     | 87.57999       | 0.54403       | 0.45597       |
| 5444392   | 11.378     | 5965                | 5726                | 1.519529    |      | 1.95393     | 1.66385     | 84.27161       | 0.51747       | 0.48253       |
| 5513861   | 11.638     | 6479                | 6411                | 1.510208    |      | 1.64494     | 1.49045     | 78.72670       | 0.56207       | 0.43793       |
| 5738698   | 11.941     | 6792                | 6773                | 4.808774    |      | 1.69030     | 1.53567     | 86.19971       | 0.55855       | 0.44415       |
| 6206751   | 12.142     | 6965                | 4885                | 1.245343    |      | 1.50527     | 1.58397     | 71.73854       | 0.77538       | 0.22462       |
| 8262223   | 12.146     | 9128                | 6849                | 1.613015    |      | 1.36443     | 1.82295     | 72.16185       | 0.58017       | 0.41983       |
| 8557388   | 12.691     | 8045                | 5328                | 1.660164    |      | 1.95352     | 1.61930     | 70.23075       | 0.91970       | 0.08030       |
| 8823397   | 13.249     | 8540                | 5724                | 1.506504    |      | 1.69962     | 1.25994     | 84.49796       | 0.85919       | 0.14081       |
| 9159301   | 12.146     | 7959                | 4209                | 3.044772    |      | 2.85734     | 1.99571     | 85.84383       | 0.95374       | 0.04626       |
| 9357275   | 12.186     | 7545                | 5580                | 1.588298    |      | 1.55017     | 1.60682     | 72.45522       | 0.80551       | 0.19449       |
| 9592855   | 12.285     | 7290                | 7087                | 1.219235    |      | 1.59043     | 1.63729     | 68.48705       | 0.49956       | 0.50044       |
| 9602595   | 11.882     | 9679                | 5705                | 3.556513    |      | 3.59066     | 81.85038     | 88.88448       | 0.31987       | 0.11156       |
| 9851944   | 11.249     | 7026                | 6902                | 2.163902    |      | 1.81528     | 2.72157     | 73.91465       | 1.30131       | 0.68013       |
| 8999416   | 10.028     | 11056               | 6278                | 1.332564    |      | 2.24077     | 1.73692     | 87.46126       | 1.88209       | 0.01194       |
| 10156664  | 10.367     | 7424                | 6268                | 4.855936    |      | 1.36860     | 1.79987     | 81.72808       | 0.55287       | 0.44713       |
| 10191056  | 10.811     | 6588                | 6455                | 2.427495    |      | 1.84616     | 1.63960     | 78.92767       | 0.56878       | 0.43122       |
| 10486425  | 12.465     | 7018                | 5847                | 2.572481    |      | 1.68109     | 0.93051     | 83.97970       | 0.93073       | 0.06927       |
| 10619109  | 11.704     | 7028                | 3903                | 2.045166    |      | 1.78472     | 2.45157     | 70.46071       | 0.84881       | 0.15191       |
| 10661783  | 9.586      | 7764                | 6001                | 1.21363     |      | 0.86273     | 2.79316     | 72.36971       | 1.18305       | 0.81695       |
| 10686876  | 11.727     | 7944                | 5842                | 2.681485    |      | 0.57829     | 2.80106     | 76.28074       | 0.21066       | 0.78934       |
| 10736223  | 13.621     | 7797                | 5069                | 1.105094    |      | 1.49741     | 1.29056     | 88.43092       | 0.88487       | 0.11513       |
| 10858720  | 10.971     | 7282                | 7223                | 0.952378    |      | 1.54288     | 1.36804     | 88.33953       | 0.57041       | 0.42959       |

Note.
$^a$ Effective temperatures are from Matson et al. (2017).
During eclipses. Pixels that were saturated have been omitted. The model systematically underestimates it. We note that the residuals are strongest that the PRF shape will change during primary and secondary eclipses. There exists a known systematic effect that may cause the PRF shape to change in a way that is correlated with the astrophysical signal, known as the “brighter–fatter” effect. This systematic effect is well documented (e.g., Antilogus et al. 2014; Guyonnet et al. 2015; Lage et al. 2017), and causes the shape of PSFs to get broader as sources become brighter, due to increased electrostatic repulsion as a pixel accumulates charge.

In the above sections, we have purposefully chosen a bright eclipsing binary with large eclipse depth. We have included terms to account for nonlinearity in each pixel, which should in theory entirely account for the brighter–fatter effect (as we are treating each pixel independently). However, the flux of the target is changing significantly, and we must verify that the shape change we have identified is not due to the brighter–fatter effect. To do this, we show the reduced chi-squared of the model for KIC 2708156 as a function of flux in Figure 9 (it is also shown in Figure 5 as a function of time).

In Figure 9, the primary and secondary eclipses have been highlighted in red and blue respectively. In reduced chi-squared, the primary and secondary eclipses clearly follow two tracks, with the secondary having a much higher reduced chi-squared than the primary at a flux of 0.9. Crucially, there are several instances where the fluxes of the primary and secondary
eclipses are the same (normalized flux ≈0.85–1), but the reduced chi-squared is significantly higher for the secondary eclipse. Despite having the same measured flux on the detector, the model residual is more significant during secondary eclipse; the pixel-level model fits more poorly during secondary eclipse, indicating a more significant shape change. Given that the measured flux is the same, the brighter–fatter effect is not able to account for this difference in PRF shape between primary and secondary eclipses, and we can discount this as the source of the PRF shape change.

3. Verification of Color Dependence Using Eclipsing Binaries

Based on Section 2.5, there is evidence that Kepler data on KIC 2708156 contain wavelength-dependent variability. While Section 2.5 contains compelling evidence that we are detecting photons at different wavelengths in each pixel, there could be further instrument systematics that we are unaware of that are causing this effect in this specific target. To verify that the effect is due to a wavelength-dependent PRF, we can use a sample of bright eclipsing binaries where we have temperature estimates for each star in the binary. Each binary has its own set of temperatures and unique orbital parameters, and as such will have a unique color variability as a function of phase.

In this section we will demonstrate the wavelength dependence of the Kepler PRF by

1. Modeling the systematics for each target, for every pixel and every quarter. We will build the same model as discussed in Section 2.2, modeling the astrophysics and variable instrument systematics in every pixel light curve.

2. Building theoretical time series for each eclipsing binary, evaluated at a coarse wavelength grid. These models will include the theoretical, expected change in eclipse depths due to the change in luminosity of each star, evaluated at a grid of wavelengths. To build these time-series models of eclipsing binaries we will assume simple blackbody emission from each star in the eclipsing binary. This time-series model of eclipsing binaries is discussed in Section 3.2.

3. Comparing the measured eclipse depths in each residual pixel time series with the anticipated theoretical light curves of eclipsing binaries at each wavelength. We will show that the residuals of the data and our pixel-level model in each pixel contain significant changes in eclipse depth (both secondary and primary), and that these residuals correlate with the anticipated wavelength variability predicted by a basic time-series model of eclipsing binaries. This comparison is discussed in Section 3.3.

3.1. Sample Selection

In this work we use a small literature sample of eclipsing binaries with temperature estimates for each component from Matson et al. (2017). From that sample, we exclude systems with nonzero eccentricities, obvious eclipse timing variations, and uncertain temperatures. Our reduced sample of 27 eclipsing binaries is given in Table 1 in the Appendix. All targets in this sample are bright Kepler eclipsing binaries with short periods. We select this small sample only as a proof of concept for the wavelength-dependent PRF, but in theory this effect would apply to all bright eclipsing binaries observed with Kepler.

3.2. Modeling Eclipsing Binaries

For each of the eclipsing binaries in Table 1 we fit the SAP light curve to find the best-fit orbital parameters (including radii and impact parameter). We use the analytic light-curve modeling code starry\(^4\) (Luger et al. 2019) to find the best-fit solution. In order to model phase-curve variation, we use simple models from Faigler & Mazeh (2011), given as

\[
\begin{align*}
\alpha_{ee} \cos(2\pi\phi + \pi) + \alpha_{el} \cos(4\pi\phi + \pi) \\
+ \alpha_{do} \sin(2\pi\phi + \pi) + c
\end{align*}
\]

where \(\phi\) is the phase of the eclipsing binary, \(\alpha_{ee}, \alpha_{el}, \) and \(\alpha_{do}\) are free parameters for the weights of the reflected, ellipsoidal, and Doppler beaming contributions respectively, and \(c\) is a

\(^4\) https://rodiluger.github.io/starry/v1.0.0/
constant. Using literature values of effective temperatures from Matson et al. (2017), and assuming a blackbody emission for each component, we then predict the time series of the eclipsing binary at 400, 500, 700, and 800 nm. Only the luminosity of each star is varied at each wavelength point, and any change in the radius of the star is assumed to be negligible.
3.3. Identifying a Trend across the Eclipsing Binary Catalog

Using this grid of wavelength-dependent light curves, we can now test whether the Kepler data on the sample of eclipsing binaries show consistent evidence of color variability, in the same way as KIC 2708156. To test this, we first build pixel model light curves for every target, for every pixel and quarter, using the same methods as discussed in Section 2.2. We then build the residuals of the true Kepler data and our data-driven pixel time-series model.

For targets where there is a strong wavelength dependence, we expect the secondary eclipse depth residual to be large. But, since each pixel may have a different effective wavelength, the residual eclipse depth varies between pixels (see the right panel of Figure 7). We define the “measured eclipse depth variation” in this data set as the standard deviation of the residual eclipse depths, measured in each pixel for observation of a single quarter, in pixels where a residual is detected at \(\geq 3\sigma\) significance. This is the standard deviation of the values in the left panel of Figure 7. We take this measurement once per quarter. In this section we choose the standard deviation as our metric for the change in eclipse depth, which allows us to combine the residuals in all pixels into a single, simple metric. (In Section 4 we use a more sophisticated approach to extract wavelength information.)

Each eclipsing binary has a unique expected color variability; binaries containing stars with similar temperatures will vary less in wavelength than stars with extreme temperatures, and binaries with deeper eclipses will have a larger relative change. To understand our sample, we need a single numeric value to assess the magnitude of the expected, theoretical wavelength dependence. We define the theoretical eclipse depth variation as the difference between the time-series models of eclipsing binaries at 500 and 700 nm, averaged during secondary eclipse. We choose wavelengths that are slightly bluer and slightly redder than the center of the Kepler bandpass. This provides a single, quantitative estimate of the magnitude of color variability anticipated for each target, based on the temperatures and orbital parameters of the star components. These wavelength values have been chosen somewhat arbitrarily to construct this simple numeric estimator, in order to demonstrate the consistent wavelength dependence in our sample of eclipsing binaries.

Figure 9. Left: the SAP flux for KIC 2708156 as a function of time. Right: the SAP flux for KIC 2708156 as a function of reduced chi-squared. The primary eclipses have been highlighted in red, and the secondary ones in blue. The primary and secondary eclipses clearly follow two tracks in reduced chi-squared, with the secondary eclipse having generally higher reduced chi-squared as a function of flux. Given that there are several instances where the fluxes from primary eclipse and from secondary eclipse are measured to be the same (normalized flux \(\approx 0.85\)), we understand the pixel-level model fit to be “worse” during secondary eclipse. As such, the PRF shape change we identify in this work cannot be attributed to simple systematic effects due to nonlinearity of flux, including the brighter–fatter effect, which would be agnostic of the “phase” of the binary.

Figure 10 shows the correlation between the theoretical eclipse depth variation and the measured eclipse depth variation, for each target. We plot the average measured eclipse depth variation across all Kepler quarters, with uncertainties given by the standard deviation of the measurement across all quarters. As shown in Figure 10, we find a statistically significant correlation between the measured variation in eclipse depth and the predicted theoretical variation across our sample. Using a Pearson correlation test, we find there is a \((\sim 3\sigma)\) correlation...
estimate the effective wavelength of each pixel. An example of this to be a nuisance parameter in this work for wavelength-dependent phase-curve variation. We therefore mask out this part of the phase curve when fitting our eclipsing binary model.

Figure 11 shows a histogram of best-fit wavelengths to every pixel in our sample, (i.e., every pixel, in every quarter, for every target). Using a simple chi-squared test, we remove pixels where the eclipsing binary model is a poorer fit than a flat line. A value of 0 indicates that there is no significant difference between the SAP flux light curve and the pixel light curve. We find the distribution of wavelengths to be approximately Gaussian, with a standard deviation of 40 nm. Figure 12 shows that, for bright eclipsing binaries observed by Kepler, a broad range of wavelengths are accessible.

3.4. Coarse Wavelength Calibration

It is beyond the scope of this initial work to fully model the PRF shape of Kepler. Fully modeling the PRF would require an investigation in the entire data set, and would provide a PRF model as a function of magnitude, spatial position on the detector, and wavelength. However, using our sample, we can provide a coarse wavelength calibration, and estimate the range of effective wavelengths for each pixel in our sample.

To provide a wavelength calibration we will fit the residuals of the Kepler data less our PRF model (Section 2) with our wavelength-dependent time-series model of eclipsing binaries, for every pixel and every quarter. We use a grid of pre-computed eclipsing binary models for each target, evaluated at 300–900 nm with a grid spacing of 100 nm. We then interpolate between these grid points to fit the best fitting wavelength model. We additionally fit phase-curve variations using Equation (9), to allow for wavelength-dependent phase-curve variation (we consider this to be a nuisance parameter in this work). Using this model we estimate the effective wavelength of each pixel. An example of this fit is shown in Figure 11 for two example pixels of KIC 2708156.

In Figure 11 it is clear that at phase 0 there is a significant, sharp change in the eclipse depth. We attribute this to wavelength-dependent changes in limb darkening. (This effect is more clearly shown in Figure 14.) Our model here is approximate, and could be improved by including physically motivated wavelength-dependent limb darkening and phase-curve variation parameters. We therefore mask out this part of the phase curve when fitting our eclipsing binary model.

Figure 12 shows a histogram of best-fit wavelengths to every pixel in our sample, (i.e., every pixel, in every quarter, for every target). Using a simple chi-squared test, we remove pixels where the eclipsing binary model is a poorer fit than a flat line. A value of 0 indicates that there is no significant difference between the SAP flux light curve and the pixel light curve. We find the distribution of wavelengths to be approximately Gaussian, with a standard deviation of 40 nm. Figure 12 shows that, for bright eclipsing binaries observed by Kepler, a broad range of wavelengths are accessible.

4. Color Aperture Photometry

As shown in the above sections, for eclipsing binaries we are able to create a coarse wavelength calibration and identify pixels that are similar in their wavelength response. For illustrative purposes, we show the “true” color of each pixel for KIC 2708156 in Figure 13. We are now able to group these pixels together and perform “color aperture photometry” (CAP). To do this, we group pixels together and create multiple “apertures” over pixels with similar measured wavelengths, across all quarters. In the context of Figure 13, we would select pixels with “similar” colors (defined by some wavelength bin) to build a color aperture. The pixels in each color aperture are then summed to create multiple time series, with one time series per aperture (wavelength bin). We can then compare these time series to show the wavelength-resolved, photometric variability of each target.

Figure 14 shows CAP light curves of four eclipsing binaries from our sample. (These targets are also highlighted in Figure 10.) Pixels were grouped into apertures in 25 nm bins, and summed over all quarters. Light curves have been grouped into panels for clarity, showing the primary and secondary eclipses, phase-curve variations, and full light curve. As
Figure 13. The “true” colors for each pixel in the first four quarters of KIC 2708156, estimated using our eclipsing binary model. White pixels are where no fit could be achieved, due to either low signal-to-noise ratio or saturation. Black pixels are (slightly) redder than the human eye can detect. This distribution is qualitatively similar to the PSF model shown in Figure 2: the “bluer” pixels are at the edges, and “redder” pixels are more concentrated toward the middle of the PRF.

Figure 14. Four example eclipsing binaries from our sample with clear variations due to the chromatic Kepler PRF. Shown are “color aperture photometry” light curves of each target, for every quarter and pixel of Kepler data, binned into the approximate wavelength bin found by this work. Each of the panels in this plot has been x- and y-scaled to show the effects of chromaticity clearly. First column: rescaled plot to show primary eclipse variations. The subtle effects of changes in limb darkening are seen in each panel. Second column: rescaled plot to show the secondary eclipse. Here the secondary eclipse depth is clearly varying across each wavelength (this is the effect used to select the wavelength bins to build the CAP light curves). It is also possible to see the subtle effects of variable limb darkening in these light curves. Third column: rescaled plot to show the phase-curve variation. In this plot it is clear that each binary has a unique phase curve, with a unique wavelength-dependent response. Fourth column: full light curve of the eclipsing binary without rescaling. The color bar shows the wavelength of each light curve, relative to the center of the Kepler bandpass. The dashed line indicates the SAP flux light curve.
shown in that figure, (1) the secondary eclipse depth varies significantly in each wavelength bin, (2) the phase curve varies in each wavelength bin, and (3) limb darkening during both primary and secondary eclipses varies. Each of these three components is expected to have a wavelength dependence, and the fact that each is independently changing is the final piece of evidence that the PRF of Kepler is wavelength-dependent, and that it is possible to extract coarse, multiwavelength photometry from Kepler data.

The CAP light curves in this work are presented only as examples, but further work could use the techniques presented here to extract multiwavelength photometry of the sample of Kepler eclipsing binaries, and estimate their fundamental properties. Kepler eclipsing binaries, and estimate their fundamental properties (such as temperature ratio, or modeling ellipsoidal variations and Doppler beaming etc.). This work could be expanded to include eclipsing binaries beyond the small sample used in this work.

Eclipsing binaries present an ideal case for CAP; it is possible to use the actual data itself to build the correct pixel apertures for each wavelength. By using the methods above, we have identified which pixels contain “similar” secondary eclipse depths (well motivated by our understanding of time series of eclipsing binaries), and summed those in individual apertures. However, many targets exist in the Kepler data set where there is no predictable simple model for anticipated color variation, and we will not be able to correctly select the CAP apertures using the target data. To build CAP apertures in such cases, we would require an in-depth investigation into the PRF across a Kepler channel, in order to identify the best CAP apertures, agnostic of the target. However, as discussed in Section 5.2, there are other uses for the model presented in this work that do not require building CAP apertures.

5. Discussion

In the Section 3 we have verified, using a sample of eclipsing binaries, that there is a clear wavelength dependence in Kepler data. The evidence presented in this work can be summarized as follows.

1. A significant PRF shape change is identified in Kepler data of eclipsing binaries, which correlates with the eclipses and phase-curve variations (Figure 5). The change in PRF shape cannot be explained by simple instrument systematics (Figure 9).

2. The PRF shape change is qualitatively similar (tight central region surrounded by a broad halo) to the expected theoretical PRF shape based on models of the chromatic aberration from the KIH (Figure 6).

3. The measured PRF shape change shows significant correlation with the theoretical, expected PRF shape change across a sample of 27 eclipsing binaries, despite each binary having unique orbital parameters, spectral types, and instrument systematics (Figure 10).

4. The secondary eclipse, primary eclipse, phase curve, and limb darkening of eclipsing binaries in our sample vary simultaneously, at different magnitudes within each eclipsing binary (Figure 14).

5. The variations in secondary eclipse depth and primary eclipse depth can be fit simultaneously across our entire sample with a well motivated eclipsing binary model, including fixed temperatures for each star component from the literature, in order to derive consistent wavelengths for each pixel that are Gaussian-distributed with a narrow wavelength range (standard deviation of 40 nm) (Figure 12).

Given this evidence, we assert that the Kepler data are weakly chromatic. In this section we discuss some limitations and possible uses of this measurable chromaticity.

5.1. Limitations

The model for the PRF we have described here is limited. In particular, we have ignored all correlations, both in time between cadences and between pixels. In reality, (1) the systematics in consecutive cadences are highly correlated, and (2) pixels that fall close to each other in the PRF are highly correlated. In addition, observations that fall on the same channel (and therefore similar pixels) of the Kepler detector should be correlated, because they experience similar systematics and exhibit similar PRF shapes. Second, instrument effects, such as undershoot, column-dependent smear, and row-dependent bias could cause significant correlation between pixels (see Van Cleve & Caldwell 2016), which we do not account for in this work for simplicity. By including all these sources of correlations, it may be possible to build a more predictive model, applicable to targets beyond bright eclipsing binaries.

Our data-driven approach fits a model to each pixel, and each new target requires a unique model. As such, using our current model, it is not possible to approximate the effective wavelength of an arbitrary pixel on the detector. By building a fully calibrated PRF model based on all the Kepler data it is theoretically possible to estimate the effective wavelength of each pixel. We leave this as further work. Using the work presented here, and a large enough sample of eclipsing binaries, it may be possible to create a model for CAP apertures, and use those apertures to build coarse color photometry from Kepler.

This work has shown that there is a color dependence in the Kepler PRF, and each pixel has a different effective wavelength. We have not explored the limits of this effect in terms of magnitude, and have restricted our study to only bright targets with extreme and predictable color variation. Further work is required to study the chromaticity of the Kepler PRF, and find the magnitude limit of the effect across the Kepler sample.

Finally, the Kepler bandpass is in fact wide, spanning 400 nm (see KIH). While we have identified the approximate effective wavelength of pixels in our sample, this is an oversimplification. In fact, each pixel should be thought of as having a unique bandpass, which is the compound effect of all of the optical systems in the telescope, and which is different to that of the average Kepler bandpass. Further investigation is needed to understand the differences between our simplistic approach and a thorough analysis that takes into account the possibility of a variable bandpass in each pixel.

5.2. Example Uses

Despite the considerations discussed in Section 5.1, and despite the limited wavelength range of the chromaticity effect, there are several ways that the community can capitalize on this coarse multiwavelength photometry. In this section we discuss some of the ways that this effect could be employed to find new insight using the archival Kepler data, including gaining a better understanding of eclipsing binaries across multiple wavelengths, vetting exoplanet candidates, identifying and
characterizing flares, and studying pulsators. There are numerous further uses for this coarse wavelength photometry that go beyond what is discussed here, including better understanding of reddening in young stars, the color dependence of supernovae, or wavelength-dependent asteroseismology measurements. This coarse multiwavelength photometry, in conjunction with Kepler’s extreme precision, can be used both for inference and for classification of objects. For example, by incorporating pixel-level modeling into machine-learning classification algorithms (e.g., Shallue & Vanderburg 2018) it may be possible to more accurately classify true transit events.

5.2.1. Classification: Exoplanet Vetting

Exoplanet candidates require vetting, in order to establish whether they are true planets or false positives (for example from blended background eclipsing binaries). The Kepler pipeline, and tools such as the Robovetter (Coughlin 2017), employed some pixel-level vetting of exoplanet candidates to aid in distinguishing true planet candidates from false positives for the Kepler mission. This vetting included analyzing centroid offsets to identify blended targets (see Twicken et al. 2020; Thompson et al. 2018). The model presented here can be used in conjunction with existing tools to vet exoplanet candidates and provide additional vetting statistics at the pixel level.

Unlike eclipsing binaries, exoplanet systems do not usually exhibit strong color dependence at visible wavelengths. Exoplanets usually do not emit in the visible, unless they have extremely high effective temperatures. As such, for true planets, we expect no changes in transit depth due to color dependence to be detectable using the techniques described in this work. We suggest that, using the model described in this work, it is possible to vet planet candidates by searching for significant shape variation in the PRF during transit. Significant PRF shape changes would indicate that the candidate planet is in fact a grazing eclipsing binary with a significant color variation. Using chromaticity of the Kepler PRF to vet planet candidates can help distinguish between grazing eclipsing binaries and grazing planets, which are otherwise difficult to distinguish (as both exhibit V-shaped transits).

Figure 15 shows an example of how the Kepler PRF chromaticity can be used to vet planet transit candidates by comparing an eclipsing binary signal to a true planet candidate. Figure 15 shows the folded light curves of true planet

---

**Figure 15.** Demonstration of how the model presented in this work can be used to identify false positive planet candidates in Kepler data. Left: folded light curve for a single quarter of Kepler data on Kepler-1b (top) and eclipsing binary KIC 8703887 (bottom). KIC 8703887 has been chosen for its similar eclipse depth and duration to exoplanet Kepler-1b, as well as similar apparent magnitude (Kepler-1b has a Kepler magnitude of 11.34 and KIC 8703887 has a Kepler magnitude of 11.05). The color bar in these panels shows the reduced chi-squared goodness-of-fit metric between the data and our pixel-level model. The eclipsing binary KIC 8703887 could be confused with a transiting planet using the time series alone. However, the reduce chi-squared shows a significant deviation from the model during transit, indicating a PRF shape change. The true planet does not show a significant shape change. We attribute this to the color variation in the light curve of the eclipsing binary. Right: residual between the data and our pixel-level model during transit, showing the significant deviation for the eclipsing binary and good agreement for the true planet. Using the reduced chi-squared metric, we can easily identify which of these targets is the true planet and which is a grazing eclipsing binary.
Kepler-1b and eclipsing binary KIC 8703887, which shows an eclipse depth consistent with a large planet. These targets have been chosen as a demonstration owing to their similar transit/eclipse depths, similar durations, and similar apparent magnitudes. In Figure 15 we show the reduced chi-squared metric of our PRF model fit for each cadence, which shows a significant disagreement during the eclipse of KIC 8703887. Kepler-1b shows no significant disagreement in reduced chi-squared during transit, showing that it is not a false positive eclipsing binary. Using the techniques presented in this work, we can identify planet candidate targets where there is significant PRF shape variation during transit, and discount them as false positives. We note that a transit/eclipse with high signal-to-noise ratio would be required to use this technique.

There are some caveats when using this method to vet exoplanet transit signals in the Kepler data set. First, it is important to understand the magnitude of the PRF shape change due to wavelength-dependent variability as a function of Kepler magnitude and location on the detector. In this work we do not identify the magnitude limit of this effect. Second, it is important to account for astrophysical effects, such as emission for hot planets and wavelength-dependent limb darkening. Accounting for such effects is beyond the scope of this work, but is enabled by the full Kepler data set.

5.2.2. Characterization: Flares

Stellar flares have been studied extensively in Kepler data (see, e.g., Davenport 2016; Yang & Liu 2019). However, flare studies with Kepler have been adversely affected by two key factors: (1) Kepler’s monochromatic bandpass makes estimates of flare energy degenerate, because there is no way to constrain the flare temperature, (2) the majority of Kepler data were taken in long-cadence mode, where each exposure is 30 minutes, which can make identification of short-term flares on timescales \(<1\) hr difficult. We propose that the chromaticity of the Kepler PRF can help alleviate both of these problems.

Figure 16 shows an example light curve of a flaring Kepler target, with the reduced chi-squared statistic of our systematics model. Flares are a significantly poorer fit than the stellar variability, showing that there is a change in the PRF shape during flares. We attribute this to a wavelength change during flares: because the flares are much hotter than the stellar photosphere we anticipate the target flux to be bluer during flares. Using the PRF shape change, it is possible to intercompare flares within a single target and investigate whether flares are all the same temperature (peak wavelength) or whether they vary.

Short-term flares on timescales of \(<1\) hr are frequently discarded by flare searches because they are indistinguishable from the signal of cosmic rays. Using the model described here it would be possible to identify true flares by using the shape of the PRF and to reject cadences where the PRF shape is more consistent with a cosmic ray than with the target PRF or the PRF shape change due to color variability. Changes in shape due to flares should be centered on the main target, and should have a shape similar to a PSF, whereas cosmic rays can appear in any pixel, and are usually either single-pixel events or sharp, bright lines. Using simple image statistics (e.g., centroiding) it would be possible to distinguish these processes. This pixel-level shape modeling could be capitalized on by machine-learning algorithms, alongside time-variable flux information, in order to better identify short-term flares in the Kepler long-cadence data set.

If a fully calibrated model for PRF shape were available, the wavelength dependence of the Kepler PRF could be used to extract temperature information on flares in the Kepler data set. Currently we do not know the parameter space that would be probed by a fully calibrated model (e.g., completeness in Kepler magnitude, flare brightness, flare temperature etc.), but in theory it will be possible to at least place limits on the flare temperature, and to intercompare different flare energies within a single target.

5.2.3. Characterization: Pulsators

Pulsating stars such as RR Lyrae are rarely studied at the precision and cadence of the Kepler spacecraft while also obtaining multiwavelength photometry. The chromatic PRF of Kepler provides a unique opportunity to study an enormous catalog of pulsators with a 4 yr long observational baseline and high-precision photometry, and to investigate the wavelength dependence of their pulsations. Figure 17 shows an example light curve of an RR Lyrae star observed by Kepler, with the reduced chi-squared statistic for our pixel-level model fit at each phase of the pulsation. The fit is significantly poorer during the rapid rise of the pulsation, indicating that there may be a significant shape change in the PRF, and therefore a color variation.
dependence, during the pulsation. Better characterization of this color change across the sample of pulsators identified in Kepler, alongside Gaia distances, may allow more robust characterization of the temperatures of these pulsators.

5.2.4. Characterization: Eclipsing Binaries

As demonstrated in this work, eclipsing binaries are some of the best targets in which to see the effect of the chromatic Kepler PSF. Eclipse depths, phase-curve variations, and limb darkening are all variable across different wavelengths. Even coarse multiwavelength photometry can enable better understanding of stellar atmospheres, when coupled with the high-cadence and long-baseline time-series photometry of Kepler. By employing a large sample of bright eclipsing binaries, and stellar spectra to obtain temperature estimates for each component, it would be possible to use this effect to understand the wavelength dependence of reflected light variation and ellipsoidal variation in eclipsing binaries, which in turn would teach us more about their outer atmospheres. It would also be possible to better calibrate limb darkening laws as a function of temperature and wavelength. Finally, with multiwavelength information, it may be possible model surface features such as spots with a wavelength-dependent component, helping to constrain spot temperatures.

5.3. Wavelength Dependence in Other Instruments and Data Sets

The effect discussed in this work is not specific to only the Kepler spacecraft. The chromaticity of the Kepler PRF is due to wavelength dependence in the optical system, namely in the PSF as it is transmitted through the optical path of the telescope, and in the detector as the PSF is recorded. Other spacecraft will have analogous effects due to wavelength dependence in their optical systems. In particular, we highlight that the TESS spacecraft is anticipated to have a similar wavelength dependence, owing to its use of lenses and large detector pixels. In fact, the magnitude of the effect for the TESS spacecraft is anticipated to be greater than that for Kepler. We encourage the investigation of other detectors to identify wavelength-dependent PSFs, which could unlock serendipitous wavelength-dependent photometry from other NASA time-series missions.

5.3.1. Modifying the Model for K2

The model discussed in this work uses a simple polynomial in the PRF centroid position to build our design matrix and to model systematics. This is a reasonable assumption, because changes in the PRF centroid are gradual (owing to focus change and velocity aberration) and are well modeled by a low-order polynomial. In order to model data from the K2 phase of the Kepler mission, this model needs to be updated. K2 data exhibit a larger, high-frequency PSF motion due to the roll of the spacecraft. This motion noise can be accounted for in our model by using the same method outlined in Vanderburg & Johnson (2014), known as the self-flat-fielding method, where centroids are converted into “arclength” (distance along the motion vector in the imaging plane) and broken up into short bins of 5–10 days. Our model can be modified to fit K2 data by removing the simple polynomials in centroid position and flux, and replacing them with a model for arclength and flux that is split into several bins of 5–10 days, in order to remove the characteristic K2 motion noise.

6. Summary

The Kepler telescope and detector have a weak wavelength-dependent response due to chromatic aberration in the optics of the telescope and due to a wavelength-dependent intrapixel sensitivity in the detector. These effects combine to give a weakly wavelength-dependent PRF, which causes a shape change in the PRF depending on the color of the target. For bright variable stars that have different flux time series at different wavelengths, this shape change can be measured in the Kepler data set.

In this work we have presented a new data-driven model for the Kepler PRF of a target, including instrument systematics, which models the effects of motion, velocity aberration, and focus change. This model does not account for small variation in PRF shape due to wavelength dependence, and so wavelength-dependent effects can be identified in the residuals between the data and our PRF model.

This work shows that for bright eclipsing binaries it is possible to extract coarse, multiwavelength photometry across a range of ≈100 nm. This has enabled us to reveal the wavelength dependence of phase curves, limb darkening, and eclipse depth for our small sample of eclipsing binaries.

In this work we have approximately calibrated the effect using literature temperature values for a small set of eclipsing binaries. This work can be improved by expanding this sample and using robust temperature estimates with uncertainties. Fully calibrating the wavelength dependence of the Kepler PRF will require a targeted effort to build a data-driven PRF model as a function of (1) target brightness, (2) target position across the detector, and (3) wavelength. While there are theoretical PRF models for the Kepler spacecraft, these do not provide the accuracy needed to resolve the wavelength-dependent effect. Given the resource provided by Gaia of a full catalog of positions in three visible color bands, it may be possible to build a data-driven PRF model by combining Gaia and Kepler data. However, further investigation is needed.

In this work we have highlighted several potential uses for this effect, ranging from classifying exoplanet targets and flare
events, to making inference using color variability for variables such as eclipsing binaries and pulsators. There are dozens of ways even coarse multiwavelength photometry can be used to enhance the legacy of Kepler. Using the model described in this work, the community will be able to access this valuable information, which has otherwise been locked in the archive.

The method of this work is to model and remove all instrument systematics using pixel-level modeling, and identify PRF shape changes at a pixel level. This method is general, and while other instruments may require a different model for pixel systematics, the methods in this work can be applied to search for wavelength dependence in the PRFs of other instruments beyond Kepler. In particular, we highlight that the TESS spacecraft observes in a similar way to the Kepler spacecraft. Due to its instrument design of using lenses and larger pixels than Kepler, TESS is expected to have an even greater wavelength dependence than Kepler. We encourage the use of the model and techniques discussed here to search for chromatic effects in the data from other spacecraft.

This work has benefited greatly from the input of several members of the Kepler team, in particular we would like to acknowledge the contributions from Douglas Caldwell and Steven Bryson, who provided excellent consultations on the PRF for Kepler. We would like to thank the reviewer of this manuscript for the excellent review, which greatly benefited this publication.

This paper includes data collected by the Kepler mission and obtained from the MAST data archive at the Space Telescope Science Institute (STScI). Funding for the Kepler mission is provided by the NASA Science Mission Directorate. STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. This research has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program. This research made use of Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013, 2018). This research made use of Lightkurve, a Python package for Kepler and TESS data analysis (Lightkurve Collaboration et al. 2018).

**ORCID iDs**

Christina Hedges © https://orcid.org/0000-0002-3385-8391
Rodrigo Luger © https://orcid.org/0000-0002-0296-3826
Jessie Dotson © https://orcid.org/0000-0003-4206-5649

Daniel Foreman-Mackey © https://orcid.org/0000-0002-9328-5652
Geert Barentsen © https://orcid.org/0000-0002-3306-3484

---

5 http://www.astropy.org

---

**References**

Aigrain, S., Hodgkin, S. T., Irwin, M. J., Lewis, J. R., & Roberts, S. J. 2015, MNRAS, 447, 2880
Antilogus, P., Astier, P., Doherty, P., Guyonnet, A., & Regnault, N. 2014, Jinst, 9, C03048
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Bedding, T. R., Mosser, B., Huber, D., et al. 2011, Natur, 471, 608
Bryson, S. T., Tenenbaum, P., Jenkins, J. M., et al. 2010, ApJL, 713, L97
Coughlin, J. L. 2017, Planet Detection Metrics: Robovetter Completeness and Effectiveness for Data Release 25, Kepler Science Document, KSCI-19114-002
Davenport, J. R. A. 2016, ApJ, 829, 23
Dimitriadis, G., Foley, R. J., Rest, A., et al. 2019, ApJL, 870, L1
Faigler, S., & Mazeh, T. 2011, MNRAS, 415, 3921
Gilliland, R. L., Chaplin, W. J., Dunham, E. W., et al. 2011, ApJS, 197, 6
Guyonnet, A., Astier, P., Antilogus, P., Regnault, N., & Doherty, P. 2015, A&A, 575, A41
Howell, S. B., Soebeck, C., Haas, M., et al. 2014, PASP, 126, 398
Jenkins, J. M. 2017, Kepler Data Processing Handbook: Overview of the Science Operations Center, Kepler Science Document, KSCI-19081-002
Jenkins, J. M., Caldwell, D. A., Chandrasekaran, H., et al. 2010a, ApJL, 713, L120
Jenkins, J. M., Caldwell, D. A., Chandrasekaran, H., et al. 2010b, ApJL, 713, L87
Lage, C., Bradshaw, A., & Tyson, J. A. 2017, JInst, 12, C03091
Lightkurve Collaboration, Cardoso, J. V. M., Hedges, C., et al. 2018, Lightkurve: Kepler and TESS Time Series Analysis in Python, Astrophysics Source Code Library, ascl:1812.013
Luger, R., Agol, E., Foreman-Mackey, D., et al. 2019, AJ, 157, 64
Luger, R., Agol, E., Kruse, E., et al. 2016, AJ, 152, 100
Luger, R., Foreman-Mackey, D., & Hogg, D. W. 2017, RNAA, 1, 7
Matson, R. A., Gies, D. R., Guo, Z., & Williams, S. J. 2017, AJ, 154, 216
O’Donovan, F. T., Charbonneau, D., Mandushev, G., et al. 2006, ApJL, 651, L61
Shallue, C. J., & Vanderburg, A. 2018, AJ, 155, 94
Silva Aguirre, V., Bojsen-Hansen, M., Slumstrup, D., et al. 2018, MNRAS, 475, 5487
Smith, J. C., Stumpe, M. C., Van Cleve, J. E., et al. 2012, PASP, 124, 1000
Thompson, S. E., Coughlin, J. L., Hoffman, K., et al. 2018, ApJS, 235, 38
Twicken, J. D., Catanzarite, J. H., Clarke, B. D., et al. 2020, PASP, 130, 064502
Van Cleve, J. E., & Caldwell, D. A. 2016, Kepler Instrument Handbook, Kepler Science Document, KSCI-19033-002
Vanderburg, A., & Johnson, J. A. 2014, PASP, 126, 948
Vorobiev, D., Ninkov, Z., Caldwell, D., & Mochnacki, S. 2018, Proc. SPIE, 10698, 106985J
Yang, H., & Liu, J. 2019, ApJS, 241, 29