

FluxCT: A Web Tool for Identifying Contaminating Flux in *Kepler* and *TESS* Target Pixel Files

JESSICA SCHONHUT-STASIK ¹ AND KEIVAN STASSUN ¹

¹*Department of Physics and Astronomy, Vanderbilt University, Nashville, TN 37235, USA**

ABSTRACT

We announce FluxCT, a web tool for identifying contaminating flux in *Kepler* and *TESS* target pixel files (TPFs) due to secondary visual sources. We demonstrate the usage of this tool and discuss the benefits of this tool over a simple *Gaia* radius search. FluxCT focuses on clarity and simplicity, where the only input needed from the user is a KIC or TIC ID. By more properly accounting for the true shape of the photometric pixel apertures, FluxCT is able to produce much more accurate estimates of contaminating flux as compared to simple radial cone searches.

1. INTRODUCTION

Contaminating flux from unexpected sources is a known issue in both *Kepler* (Borucki et al. 2010) and *TESS* (Ricker et al. 2015) light curves. Excess flux occurs due to a combination of pixel size (4'' and 21'', respectively) and flux integration — taking the sum of observed flux in all pixels comprising the Target’s Pixel Aperture (TPA). Data for each source is collected in Target Pixel Files (TPFs) and overlaid with a TPA. TPAs vary in size and shape for each source and are approximately 20'' (~ 5 pixels) in the x direction and 40'' (~ 10 pixels) in the y direction¹. The use of flux integration ensures that all light from the target is collected. Subsequently, a light curve is created by plotting the integrated flux measurements.

Any measurements determined using a light curve that contains excess flux may result in inaccurate conclusions. Some documented examples include false-positive planet transit signals caused by eclipsing binaries (Ziegler et al. 2016), the underestimation of planet radii (Ziegler et al. 2018), the possible dilution of stellar oscillations (Schonhut-Stasik et al. 2017, 2020), and observed stellar flares attributed to an incorrect source at a rate of $6.7 \pm 0.4\%$ for *Kepler* short cadence light curves (Jackman et al. 2021).

2. WEB TOOL

FluxCT is an online web tool. Use of FluxCT requires only an internet connection and access to a browser. We built FluxCT by combining Python scripts into a Python Anywhere² framework, allowing the script to run through a website. The user navigates to the site and enters the KIC or TIC ID they wish to search; FluxCT then produces the plots and data directly to the user’s browser.

Internally, once provided with an ID number, FluxCT will pull the source’s TPF data using the Python module `lightkurve` (Lightkurve Collaboration et al. 2018). Then, using the TPA information for the source, FluxCT creates a unique search polygon to contain the TPA, with as few of the surrounding pixels as possible. Next, the script searches this area using `astroquery` (Ginsburg et al. 2019) to mine the *Gaia* EDR3 database (Gaia Collaboration et al. 2021) for any potentially contaminating sources in the search polygon.

FluxCT downloads data for the potential contaminating sources from the *Gaia* database, then outputs this data, along with a plot and some calculations. Figure ?? shows an example plot for KIC 11860814. The data printed to the browser includes the magnitude of each source found, the magnitude differences between each accompanying source and the primary target, the ratio of excess flux for each accompanying source to the primary target, and the total percentage of flux in the system from all contaminants. For the flux values, we use the *Gaia* G-band flux as the closest approximation to the *Kepler* passband. FluxCT also returns the Renormalized Unit Weight Error (RUWE); often used

* NISE Fellow, Frist Center for Autism + Innovation, Nashville, TN 37212, USA

¹ Based on the average TPA size found in our test of 147 *Kepler* sources. However, TPA size will vary with the brightness of the sample, as brighter stars have larger TPAs

² <https://www.pythonanywhere.com>

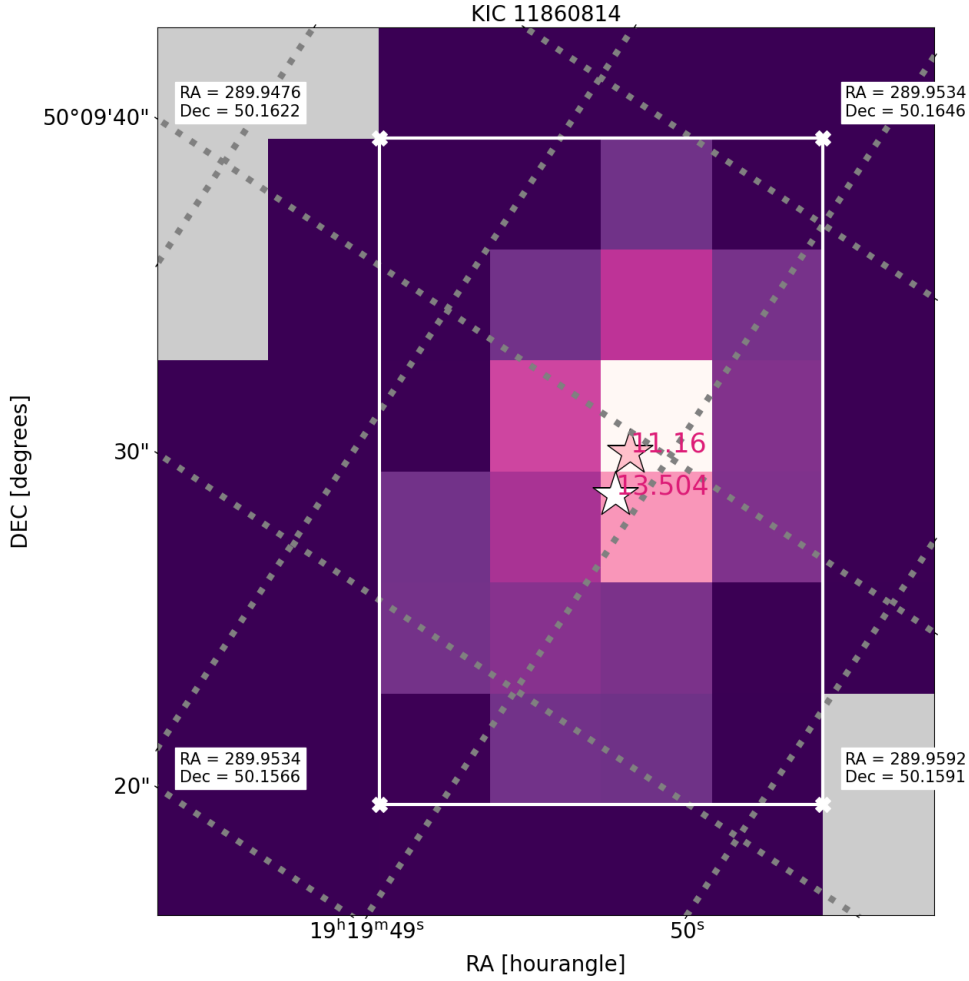


Figure 1: The web tool provides the resulting FluxCT plot for KIC 11860814. The pink star shows the target source in the RA (hourangle) Dec (degrees) space, and on-sky coordinates are shown by grey dotted lines. Each colored square depicts a pixel. Grey pixels constitute areas where no data was collected, and dark purple represents pixels not included in the TPA. Lighter shades of purple, pink, and white correspond to the TPA for this source, with the color scaled toward white as flux levels increase. The white lines around the TPA are the smallest possible polygon area encapsulating the TPA, with RA and Dec values given for each corner of the search area. KIC 11860814 has one additional source within the target pixel aperture, with a magnitude of 11.16 for the target and 13.50 for the accompanying source, giving a magnitude difference of ~ 2 . There is a flux ratio (F_2/F_1) of ~ 0.1 , resulting in a total flux contamination of 10%. The secondary source has a RUWE of 1.7 and could be an unresolved double system.

39 as a marker for possible close binarity (Lindgren et al. 2018; Stassun & Torres 2021). The use of RUWE supports
 40 completeness, as *Gaia* is only able to resolve stars down to $1''$ (Ziegler et al. 2018).

41 3. THE CASE FOR USE

42 When searching for contaminating sources, it is crucial to be mindful of the size and shape of the search area. For
 43 both *Kepler* and *TESS*, the TPA encapsulates many pixels, not just the pixel where the target source lies. Furthermore,

the target source may not fall in a pixel central to the TPA, and TPAs are not usually symmetric. Therefore, simple radial searches can prove inadequate to find the true amount of contaminating flux.

For example, in the case of *Kepler*, a common practice is to check for contaminating stellar sources using a circular 4'' radial search. The dominant issue is that the average TPA spans five by ten pixels; therefore, the small radial search will miss possible contaminants. We tested this using a sample of 147 *Kepler* stars that have observed oscillations and are currently undergoing analysis with this method to consider dilution of their oscillation amplitudes by contaminating sources (Schönhut-Stasik in prep.). In this sample, using a standard 4'' radial cone search with *Gaia* found 23 contaminant sources around the 147 stars. In contrast, FluxCT found 107 stars within *Gaia* search polygons over this sample, representing a much more complete estimate of the true contaminating flux.

Sometimes, in an attempt to find all possible sources of contamination, one may use a much larger radial search, for example, 20''. A radial search of this size risks exceeding the boundaries of the TPAs and can return many stars outside the TPA that would not add excess flux to the target system. We tested this by performing a radial search of 20'' in *Gaia*, where we found 671 sources that fall within 20'' of our 147 sources, meaning the cone search found 564 stars not identified in the above FluxCT determination—a vast overestimate. These extensive radial searches can be even more problematic for targets that fall closer to the Galactic Plane or are in areas of high stellar density.

To be sure, FluxCT is not perfect, and will itself slightly overestimate the number of contaminating sources when the FluxCT rectangular polygon is somewhat larger than the complex TPA. For the same test sample as above, a manual investigation found seven stars within the FluxCT polygon but outside the true TPA, an overestimate of 6.5%.

4. ACCESSIBILITY AND FUTURE WORK

FluxCT is currently available as a web tool that allows the user to search for single targets. The choice to initially create an accessible web tool is motivated by the desire to encourage broader accessibility to *TESS* and *Kepler* data for students. The ease of use of this tool allows students as early as high school to explore the data without needing advanced knowledge of Python. A full version that can search multiple sources is available directly from the corresponding author and will soon be available at the companion GitHub³. Any browser can access the web tool at the following link (<http://jstasik.pythonanywhere.com>), and the corresponding author welcomes any suggestions for updates. More example plots and supplemental material exist at <https://www.jessicastasik.com/flux-contamination-tool>.

This paper includes data collected by the Kepler mission and obtained from the MAST data archive at the Space Telescope Science Institute (STScI). This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>).

Software: Astropy (Astropy Collaboration et al. 2022), Astroquery (Ginsburg et al. 2019), LightKurve (Lightkurve Collaboration et al. 2018), Numpy (Harris et al. 2020) Pandas (pandas development team 2020; Wes McKinney 2010) Matplotlib (Hunter 2007)

REFERENCES

- 77 Astropy Collaboration, Price-Whelan, A. M., Lim, P. L.,
78 et al. 2022, ApJ, 935, 167, doi: [10.3847/1538-4357/ac7c74](https://doi.org/10.3847/1538-4357/ac7c74)
- 79 Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science,
80 327, 977, doi: [10.1126/science.1185402](https://doi.org/10.1126/science.1185402)
- 81 Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al.
82 2021, A&A, 649, A1, doi: [10.1051/0004-6361/202039657](https://doi.org/10.1051/0004-6361/202039657)
- 83 Ginsburg, A., Sipőcz, B. M., Brasseur, C. E., et al. 2019,
84 AJ, 157, 98, doi: [10.3847/1538-3881/aafc33](https://doi.org/10.3847/1538-3881/aafc33)
- 85 Harris, C. R., Millman, K. J., van der Walt, S. J., et al.
86 2020, Nature, 585, 357, doi: [10.1038/s41586-020-2649-2](https://doi.org/10.1038/s41586-020-2649-2)
- 87 Hunter, J. D. 2007, Computing in Science & Engineering, 9,
88 90, doi: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55)
- 89 Jackman, J. A. G., Shkolnik, E., & Loyd, R. O. P. 2021,
90 MNRAS, 502, 2033, doi: [10.1093/mnras/stab166](https://doi.org/10.1093/mnras/stab166)
- 91 Lightkurve Collaboration, Cardoso, J. V. d. M., Hedges, C.,
92 et al. 2018, Lightkurve: Kepler and TESS time series
93 analysis in Python, Astrophysics Source Code Library.
94 <http://ascl.net/1812.013>
- 95 Lindgren, L., Hernández, J., Bombrun, A., et al. 2018,
96 A&A, 616, A2, doi: [10.1051/0004-6361/201832727](https://doi.org/10.1051/0004-6361/201832727)
- 97 pandas development team, T. 2020, pandas-dev/pandas:
98 Pandas, latest, Zenodo, doi: [10.5281/zenodo.3509134](https://doi.org/10.5281/zenodo.3509134)

³ <https://github.com/Jesstella/FluxCT>

- 99 Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015,
100 Journal of Astronomical Telescopes, Instruments, and
101 Systems, 1, 014003, doi: [10.1117/1.JATIS.1.1.014003](https://doi.org/10.1117/1.JATIS.1.1.014003)
- 102 Schonhut-Stasik, J., Huber, D., Baranec, C., et al. 2020,
103 ApJ, 888, 34, doi: [10.3847/1538-4357/ab50c3](https://doi.org/10.3847/1538-4357/ab50c3)
- 104 Schonhut-Stasik, J. S., Baranec, C., Huber, D., et al. 2017,
105 ApJ, 847, 97, doi: [10.3847/1538-4357/aa886f](https://doi.org/10.3847/1538-4357/aa886f)
- 106 Stassun, K. G., & Torres, G. 2021, ApJL, 907, L33,
107 doi: [10.3847/2041-8213/abdaad](https://doi.org/10.3847/2041-8213/abdaad)
- 108 Wes McKinney. 2010, in Proceedings of the 9th Python in
109 Science Conference, ed. Stéfan van der Walt & Jarrod
110 Millman, 56 – 61, doi: [10.25080/Majora-92bf1922-00a](https://doi.org/10.25080/Majora-92bf1922-00a)
- 111 Ziegler, C., Law, N. M., Baranec, C., et al. 2016, in Society
112 of Photo-Optical Instrumentation Engineers (SPIE)
113 Conference Series, Vol. 9909, Adaptive Optics Systems V,
114 ed. E. Marchetti, L. M. Close, & J.-P. Véran, 99095U,
115 doi: [10.1117/12.2231185](https://doi.org/10.1117/12.2231185)
- 116 Ziegler, C., Law, N. M., Baranec, C., et al. 2018, AJ, 156,
117 83, doi: [10.3847/1538-3881/aace59](https://doi.org/10.3847/1538-3881/aace59)