April 30, 2023
Typeset using LaTeX **twocolumn** style in AASTeX631

Searching for Eclipsing Binaries in the GSFC-Eleanor-Lite Database

Andy Teng[1]

[1]*Seven Lakes High School*

ABSTRACT

This paper presents the discovery of 66 new eclipsing binaries from a cursory search of sector 3 of the GSFC-eleanor-lite database using a 2-algorithm data pipeline. This database, which contains eleanor-processed full-frame image light curves, was analyzed with a pipeline consisting of two algorithms: ECLIPSR and BLS. Within the paper, I present details on the pipeline and the methods used to verify the existence of the eclipsing binaries. I also provide a catalog of the 66 new eclipsing binaries, including the properties measured. In contrast with most other papers, this pipeline is tailored toward running on a home computer. It consists of per-light curve runtimes of approximately one second while delivering marginally less accurate detections than other approaches. Finally, this work also identifies a lower eclipsing binary density in the GSFC-eleanor-lite database compared to most short cadence photometry analyses but simultaneously possesses a higher percentage of undiscovered eclipsing binaries.

Keywords: TESS, Eclipsing Binaries

1. INTRODUCTION

The Transiting Exoplanet Survey Satellite (TESS) is a currently operational NASA Explorer-program spacecraft in its second mission extension and has been in service for a current total of five years (Barclay, 2023).

The satellite's main instruments are a set of four identical wide field charged-coupled device (CCD) cameras, with each camera having a field of view of 24° × 24° for a total of 24° × 96° of vision. TESS cameras observe continuously in a stream of two-second exposures and are summed over different periods of time for its different cadence types. For TESS' short cadence photometry, images of selected targets were created in two minute intervals. This short cadence imagery has already proven to be extremely beneficial for the detection of exoplanets and other selected deep space object types over the course of its mission (Olmschenk et al., 2021). Additionally, long cadence photometry was taken once every 30 minutes during the primary mission period and stored in Full Frame Images (FFIs) (Barclay, 2017). Later, FFI cadences were reduced to 10 minutes and then 200 seconds for the first and second mission extensions, respectively. FFIs combine the uncorrected pixel data from the entire camera. Thus, these FFIs not only include planetary transits, but also have the potential to store data on eclipsing binaries (Kruse et al., 2019).

Eclipsing binaries (EBs) are a type of binary star system orbiting within the plane of the Earth in which the two stars periodically pass in front of each other, causing a temporary decrease in brightness that can be observed from Earth. It is this decrease in brightness (transits) that all detection methods examine to determine whether the target is an EB, exoplanet, or another object entirely. In order to apply such detection methods, astronomers measure and plot the amount of received light over time, in what are called light curves. For an EB, the different sizes and luminosities of the two component stars create pronounced odd and even transits which can be analyzed (for an example of such a light curve, see Fig. 3).

Through this analysis, EBs provide valuable insights into the physical properties of stars, such as their masses, radii, and temperatures, which can be difficult to measure directly (Prsa et al., 2022). Astronomers can determine the size, shape, and periods of the stars' orbits and thus their masses through Newton's laws of gravitation; additionally, the duration and depth of the eclipses can be used to derive the stars' physical properties such as through the Stefan-Boltzmann law. Eclipsing binaries have played a crucial role in the development of modern astrophysics, providing a way to test theoretical models of stellar structure and evolution, and enabling astronomers to explore the relationships between stellar parameters. Furthermore, they are important tools for determining the distances to other galaxies by measur-

ing the spectra, which helps to refine our understanding of the scale and structure of the universe. Finally, EBs also serve as an important safeguard against false positives for exoplanet detection, as the light curves of the two objects can appear to be similar due to the transit shape mostly being identical (Prsa et al., 2022). Thus, EBs were chosen as the target to examine FFI data over other transient phenomena for their lack of study within the literature in comparison with their importance.

2. LITERATURE REVIEW

As the name of the spacecraft would suggest, the analysis of research data from TESS over the past few years has mainly focused on the discovery of exoplanet candidates. Although this paper focuses on eclipsing binaries (EBs), the same issues that plague exoplanet transit candidates affect EB detection because both use transits as an identification method. The primary difference between the two is the different transit depths and durations for odd/even transits in EBs, as mentioned previously. Only one final step of comparing odd and even transits during vetting is necessary to distinguish the two. Thus, it is necessary to examine the methodology that past researchers have used to identify both exoplanets and EBs in order to adequately understand the current situation for EB detection. To begin, it is necessary to investigate the fundamental principles that underlie such detection methods. Early on in exoplanet discovery, observation depended on Earth-based telescopes performing radial velocity analysis, where the gravitational effect of planets or other gravitational bodies on the parent star induced changes in its rotational velocity which can be viewed from Earth in the form of redshift in its spectrum (Mayor & Queloz, 1995). However, beginning with Kepler (Borucki et al., 2007) and later TESS (Ricker et al., 2014), astronomers could use much more refined photometry from orbital telescopes to identify the physical crossing of a planet in front of a star. Light curves gathered could then be used to find transit candidates.

Although such light curves produced from FFIs are a treasure trove for discoveries, the sheer volume of data is time and power-consuming to process. As such, being able to efficiently extract and process light curves is crucial for a successful algorithm. One common method of doing so was developed by Feinstein et al. (2019), who presented a new process of extracting light curves from FFIs, called *eleanor*. This Python package includes functionality that reduces noise and streamlines light curve analysis, including a point spread function (PSF) and primary component analysis (PCA) to create a corrected flux for each light curve generated. However, for some of the sectors, Feinstein et al. note that artifacts and background stars/planets would taint the data. Thus, it is important to find light curves from the FFIs of non-disrupted sectors and/or from time periods where such artifacts are not present.

In order to perform analysis of these light curves a variety of techniques are used, including the naive but standard box least squares (BLS) algorithm (Kovacs 2002, 2016), a multitude of machined learning approaches (Osborn et al. 2019) (Ofman et al. 2021), and Bayesian probability (Giacalone et al. 2021). Although all of these algorithms have been demonstrated to have their use cases in determining EBs and exoplanets, they each have varying drawbacks that must be considered. In general, machine learning is highly versatile but suffers from long training times and needing large, preclassified data sets to train on. On the other hand, simpler algorithms such as BLS are faster to run but are generally less accurate. BLS applies a strict-period, box-fitting search that guesses and checks its way through transit durations and depths with a periodogram (Kovacs 2002). In between these two extremes, are the statistical approaches, which can range from the simple Z-score checking and autocorrelation function (ACF) of Howard et al. (2022) to the complex, Bayesian modeling of Giacalone et al. (2021). It is important to balance both effectiveness and efficiency: an algorithm must both have high accuracy for a given inputted light curve while also being time-efficient to run.

One such example of an optimized algorithm that stands out from the previously mentioned ones is by Kruse et al. (2019), who modified the Quasi-automated Transit Search algorithm (QATS) (Carter and Agol 2013) to classify transit data from the K2 mission. QATS is more flexible than the strict periodicity of the BLS algorithm and instead only examines consecutive transits that fall into a time window. Thus, QATS can detect exoplanet candidates and EBs that exhibit transit timing variance (TTV) while being simpler than the more complex approaches by only needing to examine the transits within a few specific windows rather than the entire light curve. After the QATS algorithm was used on the data, human vetting was needed to complete the process of filtering out transit candidates, as is standard in the field. Although the QATS algorithm has strong accuracy and is relatively efficient, it still maintains some drawbacks due to the fact that it was designed specifically for Kepler exoplanets. Additionally, my data pipeline was exclusively coded in Python for the sake of simplicity and time efficiency, and thus it was decided that the C++ coded QATS would not be included.

Due to the varying drawbacks and benefits of each algorithm, it is common to combine them into what are called data pipelines. One recent pipeline, created by Olmschenk et al. (2021), presented a neural network model for identifying planetary transit candidates in TESS FFIs. Their convolutional neural network (CNN) runs on the Keras and Tensorflow machine learning python packages. Data is fed in through several preprocessing stages, including *eleanor* (Feinstein et al., 2019) as well as QATS (Kruse et al., 2019). Finally, vetting was done with the Discovery and Vetting of Exoplanets (DAVE) (Kostov et al. 2019). The CNN is a 1-dimensional deep neural network with 11-13 layers that can perform fitting on a light curve in reportedly $\sim$ 5 ms. The CNN, dubbed *ramjet*, was used to determine possible planet transit candidates and then classify them based on the process it learned during its training phase, which consisted of confirmed ExoFOP-TESS-confirmed planets. Even though it was possible for the CNN to identify EB light curves within the FFIs as well, the primary focus of the training set was exoplanets, and thus most outputted EB candidates were rather rejected exoplanets, demonstrating a lack of focus on EBs that was exemplary of most reviewed sources. The overall success rate on their testing data set reached nearly 90% and was generally time-efficient during training compared with implementations of machine learning approaches. However, their data pipeline still required a vast amount of resources to run and train the CNN which would be impossible for me to replicate, mainly due to time constraints. Clearly, a theme of overly complex algorithms combining to form computationally expensive data pipelines is beginning to emerge.

Indeed, this pattern continues in the few EB-focused papers published. Specifically, the most significant piece of literature published on TESS EBs, written by Prsa et al. (2022), examined short cadence light curves with a 3-algorithm data pipeline followed up by vetting with DAVE and modeling. They examined approximately 200,000 targets selected from several preexisting public catalogs and databases and subsequently generated a list of 4,584 EBs from TESS sectors 1-26 to be added to an initial catalog. This catalog was later expanded by 370 new candidates identified from short cadence light curves by Howard et al. using another 3-stage pipeline with vetting. In order to create statistical thresholds to differentiate EBs that pass through the algorithm, they manually classified several thousand light curves, ultimately arriving at a multi-pass subtractive smoothing method and creating several heuristic thresholds. On the initial pass, a maximum power of greater than 1,500 for the periodogram was determined, which would signify the presence of an EB within that light curve. The pipelines of the two previously discussed papers were thorough, but confirms two gaps within EB identification literature. First, the majority of existing data pipelines are too complex or demanding on hardware for budget-limited researchers to implement, with three algorithms or multiple passes being done per light curve. Second, even within the already limited EB literature, there still exists a focus on short cadence photometry as opposed to FFIs.

To summarize, it is evident that the majority of previously mentioned literature have mainly centered on exoplanet identification, and thus there appears to be a research gap in EB identification (Prsa et al. 2021). Additionally, upon closer examination of the limited number of EB-oriented sources, there is a significant lack of analysis of FFIs, with almost all targeting short cadence light curves for their relative ease of access. With the context of prior data pipelines in place, the following methodology was designed to investigate the efficacy of a novel FFI database, GSFC-eleanor-lite, and suggest possible candidates for new EBs using a new, lightweight data pipeline.

3. METHODOLOGY

3.1. *Dataset*

The GSFC-eleanor-lite dataset was chosen to take advantage of both the aforementioned eleanor FFIs and access to the most recent data. In the past, the short cadence photometry, which singled out specific targets, was significantly easier to examine for research; at the time of data pipeline creation, the most recent data release was for Sector 3, which coincidentally had the greatest number of light curves of any of the released sectors. As such, it was chosen to provide a diverse range of light curves to select from. However, searching through all 1.9 million light curves would be unfeasible, so a subset of 50,000 light curves was chosen from the database. For this subset of 50,000, the data pipeline was run to demonstrate its faster runtime compared to previous pipelines and the usefulness of the GSFC-eleanor-lite database as a source of EBs and exoplanet candidates. The CSV of the dataset entries was first downloaded from the MAST portal, and then processed down to only nine digit Tess Input Catalog (TIC) numbers and those with the requisite quality flags using Python, arriving at a final dataset of 47,298 light curves. After that, FITS files of the light curves, stored in the eleanor-lite format, were queried from the Mikulski Archive for Space Telescopes using the astroquery Python package and was then inputted into the data pipeline itself. The eleanor-lite format discards all non-

essential data after processing with *eleanor* itself, enabling easier bulk downloads (Powell et al., 2022).

3.2. *Overview*

The methodology of this paper consists of a secondary data analysis performed with a data pipeline made up of two primary detection algorithms. The two, arranged in order of operation within the data pipeline, are the Eclipse Candidates in Light curves and Inference of Period at a Speedy Rate (ECLIPSR) algorithm (IJspeert et al. 2021) and the box least squares (BLS) algorithm (Kovacs 2016).

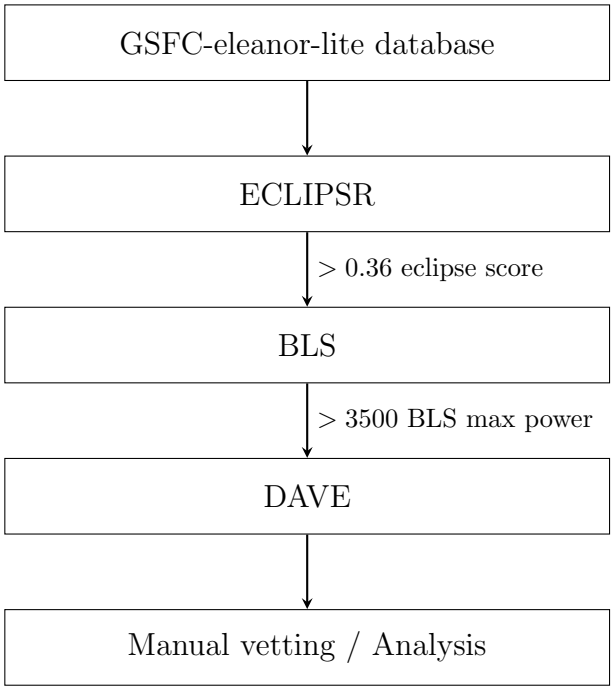
First, (paraphrased from IJspeert et al., who created the algorithm) ECLIPSR functions by determining transit locations and then finding periodicity, with the algorithm distinguishing between flat-bottomed, V-shaped, and full eclipses. With these three light curve models, it uses multiple derivatives of the light curve to detect the slopes occurring at the beginning and end of eclipses. By multiplying the first and third derivatives, the signal is amplified while noise is smoothed out. The peaks in the second derivative indicate the start and end points of the eclipse as well as marking the flat bottom of an eclipse. Additionally, the signal to noise ratio (SNR) is measured to remove false peaks that appear as transits, but are in actuality just noise. After the correct peaks are identified as transits, ECLIPSR then starts period-searching. A pre-determined range of periods is then scanned, and the highest goodness-of-fit is used to determine the best period. The range of periods is determined by the minimum and median separation between eclipse candidates, and it is checked whether doubling the period results in distinguishable secondary eclipses (IJspeert et al., 2021). If not, the algorithm checks for eclipses that have the same period but are at a different phase offset. Once the period has been determined, the light curve is assigned a score based on how closely it matches the model; a score of more than 0.36 is considered a proven EB as found by IJspeert et al.. This algorithm was chosen for its simplicity and repeatability, with only one transit necessary to fit the model and confirm the existence of an EB.

Although the previous algorithm is usually more than sufficient to identify the light curves necessary for later model fitting and triage, it is always beneficial to do a final run through of a relatively simple algorithm for final confirmation. BLS is significantly more time-consuming than ECLIPSR, so it is important to run this algorithm last to reduce the amount of false positives that erroneously run through it. One of the main advantages of BLS over other "simple" algorithms is that its transit model consists entirely of background with strict, periodic rectangular transits. Thus, with so few parameters that need to be tuned for each pass, the algorithm remains one of the fastest of the naive approaches (Kovacs 2002). Using BLS, accurate half-periods, periods, and double-periods can be generated to determine the presence of an EB over an exoplanet, and guarantees its detection if it falls within the tested range. For this paper, detached EBs with periods between 0.5 and 27 days were targeted due to this range of periods including the majority of known EBs within TESS data (Prsa et al. 2022).

The data pipeline was run in multiple stages to determine the efficacy of the various constituent algorithms and to ensure that everything is running properly. First, a dry run of the first thousand light curves is run through with only BLS to determine a baseline efficacy of the data pipeline and to determine the time needed to download and process the files. Then, a more thorough run of 10,000 light curves is performed with both BLS and ECLIPSR to better understand the full data pipeline and to characterize its speed. Finally, the final run of the full dataset is performed for characterization of the database itself. With this final run through, the final list of candidates is generated, which was then vetted and compared against the existing Prsa et al. database. After vetting is done, it is possible to perform modeling or other analysis on the light curves, such as through the Python package ELISa (Cokina 2021), which runs an affine-invariant Markov-chain Monte Carlo simulation to generate stellar parameters; however, that is mostly outside of the scope of this paper.

Candidates that are generated from the final, full run of the data pipeline are subsequently passed through DAVE and subjected to visual examination to remove false positives. DAVE is an industry standard Python program that applies centroid and flux analysis to determine whether inputted candidate light curves are valid. For visual examination, if evident primary and secondary transits could be confirmed to be periodic, they were classified as confident. If transits were ambiguous due to noise or not being differentiable between primary and secondary transits, then they were subsequently classified as probable. Light curves were determined to be non-EBs when the transit was found to be bad data. This appeared as a drastic angular decrease in flux, oftentimes going below zero in the normalized flux graph as opposed to a less than ten percent decrease in received flux for true transits. Finally, all accepted light curves that passed vetting were added to a file and compared against the existing EB database to test the validity of this data pipeline and to find if any new EB candidates were generated from this process.

The use of two, relatively fast algorithms within this data pipeline, as opposed the norm of three-algorithms within the existing literature, was hypothesized to be more efficient and result in less false negatives than past approaches. On the other hand, it would allow more false positives into the final candidates list and the added time needed for manual vetting may significantly impact the efficacy of the pipeline. For the GSFC-eleanor-lite database itself, the non-discriminatory nature of the light sources contained within the FFIs resulted in the belief that it would contain a lower, but still substantial EB density than short cadence light curve data. This factor could also artificially inflate the false positive rate beyond what the data pipeline itself would normally output.



4. RESULTS

During the first dry run of only the BLS algorithm, 1,000 light curves were analyzed, yielding 52 candidates in 2 hours. This result is significantly higher than the 0.5% confident EB rate of Howard et al. However, after close examination of the light curves, it was clear that manual vetting was needed to validate candidates. For the initial run of the data pipeline with only the BLS algorithm, there were a large number of false positives that were not caught by the data quality flags of the database. After manual vetting, only 14 of the original 52 candidates proved to be actual EBs, which is more in line with Howard et al.'s findings.

During the second run through, going through the first 10,000 light curves of the dataset, the original candidates list was revised and added to. This second run consisted of both ECLIPSR and BLS, with extracted periods being matched to ensure that the two algorithms were in agreement. The pipeline generated 273 light curves from the 10,000 processed. It is important to point out that light curves that triggered bad data quality flags or threw errors were immediately rejected and thus future investigation is necessary to determine if they are actual EBs. However, the fact that past researchers flagged it during initial processing makes this concern unlikely. After analyzing the runtime of the various algorithms, it was noted that a large portion of the time processing each light curve was dedicated to downloading the file itself (up to 10 seconds in some cases), and processing of most light curves through both algorithms took only ~ 1 second.

For the final run of the total 47,298 light curves, 1,353 candidates were generated. The previous estimate of ~ 1 second per light curve for the data pipeline was further confirmed, with median times of ~ 0.1 second for ECLIPSR and ~ 0.9 second for BLS. After this final pipeline run, analysis began with visual confirmation of candidate light curves with a custom Python program to manually classify candidates into the following three categories as described in the methodology: non-EBs, probable EBs, and confident EBs. Examples of these three classifications are provided below as reference for manual vetting procedures. From the 1,353 initial candidates after DAVE, further analysis and vetting confined the list to 70 confident and 40 probable EBs, taking approximately one and a quarter hours to complete. The list of candidates was thoroughly examined to remove false positives. Additionally, EBs that shared the same period were only counted once, as these TICs most likely belonged to the same object. If such TICs were included, there would be 124 final candidates.

After comparing with the existing Prsa et al. database, six of the candidates were present.

Figure 1. Non-EB, transit is bad data

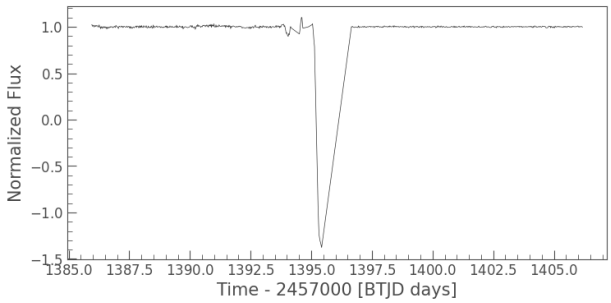


Figure 2. Probable EB, ambiguous odd/even transit, not enough transits, or too noisy

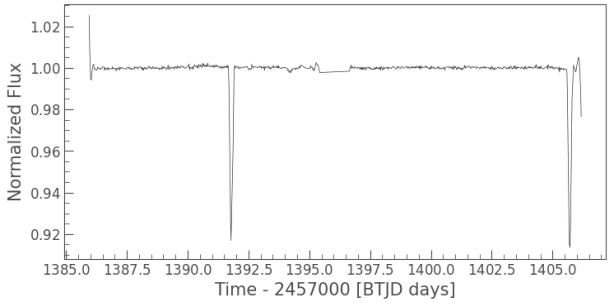
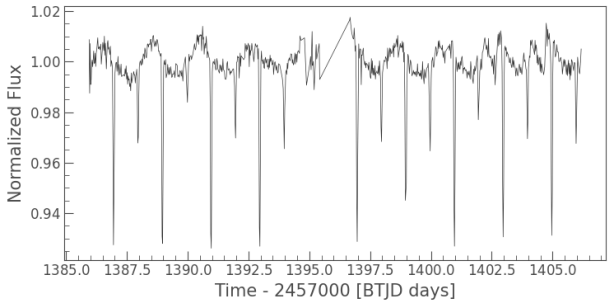


Figure 3. Confident EB, note the obvious primary and secondary transits



5. DISCUSSION

Although the success rate is relatively low (and thus the false positive rate is relatively high) compared to past studies, such as Howard et al. with a 40-50% success rate compared to only approximately 9% for this pipeline, this was the hypothesized outcome, albeit lower than expected. The faster runtime and less stringent conditions of the algorithms reduced the risk of false negatives. However, this benefit was offset by the time required to perform manual vetting, which ended up being significantly more than expected. For future studies, the pipeline may need to be tuned to decrease the false positive rate and to verify whether it is the pipeline, the database, or a combination of both that is causing this result.

For the total dataset, 124 of the 47,298 light curves were identified as probable or confident EB candidates, resulting in an EB pass rate of $\sim$ 0.26% (0.23% excluding duplicate objects). This relatively low rate is partially explained by the methodology of the original GSFC-eleanor-lite database, which processed all FFI light curves through *eleanor* with a TESS magnitude of greater than 16, thus increasing the number of included light sources that do not contain useful data. Additionally, the low pass rate is also partially caused by the momentum dump and data downlink periods at the apogee of each orbit, which creates gaps in the flux that are in some cases interpolated in the data as transits. This apparent transit would usually be caught by ECLIPSR, which is capable of handling noise, but in rare circumstances would take on the exact shape of a transit and thus pass through both algorithms as BLS does not take into account transit shape. Despite the goal of eventually reducing the appearance of false positives in general, manual vetting remains an essential step in ensuring the validity of exoplanet and EB candidates such as by citizen science projects like Planet Hunters (Eisner et al., 2021). Further studies should be conducted to determine the effectiveness of other automated methods in identifying true positives and minimizing false positives as well as reducing the necessity of projects like Planet Hunters.

Prsa et al. and later Howard et al. confirmed an approximate EB density within short cadence light curves of 0.5%. In comparison with this paper's 0.23%, the density is lower as expected and explained earlier, but the true value of the database comes from the fact that most of the final candidates were not present in the existing catalog, highlighting the importance of FFIs to future EB research. Howard et al.'s follow up on Prsa et al. yielded 370 new additions to the catalog out of roughly 2,500 final candidates, approximately 15%. After analyzing the final candidates by checking with the existing catalog, only six of the final 110 were already present, with four being identified within the pipeline's candidates as confident EBs. Thus, from the list of confident candidates, a total of 66 new known EBs was generated from the data pipeline, roughly half of all candidates and significantly greater than the 15% of Howard et al. (see appendix). When compared with previous short cadence efforts, the density of *new* EBs is significantly higher for GSFC-eleanor-lite, with a 0.22% (104 out of 47,298) rate compared to a 0.07% rate for Howard et al. (370 new EBs out of approximately 500,000 initial light curves).

Furthermore, there were only seven EBs present within the preexisting catalog not found by the data pipeline. Out of these seven, five were extremely short period EBs outside of the 0.5 to 27 day period range set for the scope of the data pipeline. After rerunning the pipeline with an updated period range on the five new light curves, they each passed without issue. Another light curve of the seven contained in the catalog but outside the candidates list appeared to be either an error on the part of the original entrant or had a period longer than the length of the sector 3 imaging time as no transits were seen for the entirety of the light curve. The final missed EB was missed likely due to a large data gap in the center. Thus, the new data pipeline performed as

expected, with extremely few false negatives and a large number of false positives, which future research would likely seek to rectify. Additionally, the results demonstrate that even with a reduced, 2-algorithm, single-pass pipeline, success rates are not diminished much, if at all. With its low computational requirements and efficiency improvements over past approaches, the data pipeline has proven itself as an effective tool for EB determination in FFI databases.

5.1. *Limitations*

The primary constraint of this approach is the time required to run the programs, which limits the size of the dataset that the algorithms can be performed on. The program was run on approximately 50,000 light curves of the several million contained within the GSFC-eleanor-lite Sector 3 database. With a limited dataset, the EB density calculated from this data pipeline may be less generalizable. Future research is necessary to fully characterize the efficacy of GSFC-eleanor-lite without such restrictions and to cut down on the issue of longer manual identification times, but nevertheless, the initial findings do show extreme promise. Furthermore, not an insignificant number of target TICs were identified as the same object, which arises from the nature of the GSFC-eleanor-lite database itself. However, it is possible this was not a major issue as the occurrence of duplicate EBs also likely meant the duplication of non-EBs. Thus, future research could attempt to determine if such occurrences actually present an issue and if so, rectify it. For the database, future papers may be necessary to examine other sectors and a wider range of targets to further determine its effectiveness as a source of EBs, in addition to other astrophysical phenomena. This paper only analyzed a small portion of the database as an initial characterization to determine its potential.

6. CONCLUSION

This research addresses the existing gap in the literature by examining a new dataset: the GSFC-eleanor-lite database (Powell et al., 2022). The new data pipeline, consisting of ECLIPSR and BLS, has been proven to both be fast and effective for targeted EBs. Furthermore, it was found that the GSFC-eleanor-lite database has a lower EB density than that of the short cadence photometry analyzed by Howard et al. but is an effective source of new EBs as opposed to mostly confirming past data. As such, all research goals initially set forth were accomplished.

Future research can seek to confirm or deny the initial insights provided by this simplified data pipeline as well as further determine if the database is worth mining as a source of EBs. As a result of this research, 66 new EBs were discovered and several intriguing directions for future study were revealed.

Software: Astroquery (Ginsburg et al., 2019), eleanor (Feinstein et al., 2019) IPython (Perez & Granger, 2007), Lightkurve (Lightkurve Collaboration, 2018), NumPy (Oliphant, 2007), Matplotlib (Hunter, 2007), Python

REFERENCES

[1] Barclay, T. (2017). *Data Products*. TESS; NASA - TESS Science Support Center. https://heasarc.gsfc.nasa.gov/docs/tess/data-products.html

[2] Barclay, T. (2023). *The Second Extended Mission*. TESS; NASA - TESS Science Support Center. https://heasarc.gsfc.nasa.gov/docs/tess/second-extended.html

[3] Borucki, W., Koch, D., Basri, G., Batalha, N., Brown, T., Caldwell, D., Christensen-Dalsgaard, J., Cochran, W., Dunham, E., Gautier, T. N., Geary, J., Gilliland, R., Jenkins, J., Kondo, Y., Latham, D., Lissauer, J. J., & Monet, D. (2007). Finding Earth-size planets in the habitable zone: the *Kepler* Mission. *Proceedings of the International Astronomical Union*, 3(S249), 17–24. https://doi.org/10.1017/s174392130801630x

[4] Carter, J. A., & Agol, E. (2013). THE QUASIPERIODIC AUTOMATED TRANSIT SEARCH ALGORITHM. *The Astrophysical Journal*, 765(2), 132. https://doi.org/10.1088/0004-637x/765/2/132

[5] Cokina, M., Fedurco, M., & Parimucha, Š. (2021). ELISa: A new tool for fast modelling of eclipsing binaries. *Astronomy & Astrophysics*, 652, A156. https://doi.org/10.1051/0004-6361/202039171

[6] Eisner, N. L., Barragán, O., Lintott, C., Aigrain, S., Nicholson, B., Boyajian, T. S., Howell, S., Johnston, C., Lakeland, B., Miller, G., McMaster, A., Parviainen, H., Safron, E. J., Schwamb, M. E., Trouille, L., Vaughan, S., Zicher, N., Allen, C., Allen, S., & Bouslog, M. (2020). Planet Hunters TESS II: findings from the first two years of TESS. *Monthly Notices of the Royal Astronomical Society*, 501(4), 4669–4690. https://doi.org/10.1093/mnras/staa3739

[7] Fausnaugh, M. M., Vallely, P. J., Kochanek, C. S., Shappee, B. J., Stanek, K. Z., Tucker, M. A., Ricker, G. R., Vanderspek, R., Latham, D. W., Seager, S., Winn, J. N., Jenkins, J. M., Berta-Thompson, Z. K., Daylan, T., Doty, J. P., Fűrész, G., Levine, A. M., Morris, R., Pál, A., & Sha, L. (2021). Early-time Light Curves of Type Ia Supernovae Observed with TESS. *The Astrophysical Journal*, 908(1), 51. https://doi.org/10.3847/1538-4357/abcd42

[8] Feinstein, A. D., Montet, B. T., Foreman-Mackey, D., Bedell, M. E., Saunders, N., Bean, J. L., Christiansen, J. L., Hedges, C., Luger, R., Scolnic, D., & Cardoso, J. V. (2019). Eleanor: An open-source tool for extracting light curves from the TESS full-frame images. *Publications of the Astronomical Society of the Pacific*, 131(1003), 094502. https://doi.org/10.1088/1538-3873/ab291c

[9] Giacalone, S., Dressing, C. D., Eric, Collins, K., Ricker, G. R., R. Vanderspek, Seager, S., Winn, J. N., Jenkins, J. M., Barclay, T., Khalid Barkaoui, Cadieux, C., Charbonneau, D., Collins, K. A., Conti, D. M., Doyon, R., Evans, P., Mourad Ghachoui, Gillon, M., & Guerrero, N. (2021). Vetting of 384 TESS Objects of Interest with TRICERATOPS and Statistical Validation of 12 Planet Candidates. *The Astronomical Journal, 161*(1), 24–24. https://doi.org/10.3847/1538-3881/abc6af

[10] Ginsburg, A., Sipőcz, B. M., Brasseur, C. E., Cowperthwaite, P. S., Craig, M. W., Deil, C., Guillochon, J., Guzman, G., Liedtke, S., Lim, P. L., Lockhart, K. E., Mommert, M., Morris, B. M., Norman, H., Parikh, M., Persson, M. V., Robitaille, T. P., Segovia, J.-C., Singer, L. P., & Tollerud, E. J. (2019). astroquery: An Astronomical Web-querying Package in Python. *The Astronomical Journal, 157*(3), 98. https://doi.org/10.3847/1538-3881/aafc33

[11] Hippke, M. & Heller, R. (2019). Optimized transit detection algorithm to search for periodic transits of small planets. *Astronomy & Astrophysics*, 623. https://doi.org/10.1051/0004-6361/201834672

[12] Howard, E. L., Davenport, J. R. A., & Covey, K. R. (2022). 370 New Eclipsing Binary Candidates from TESS Sectors 1–26. Research Notes of the AAS, 6(5), 96. https://doi.org/10.3847/2515-5172/ac6e42

[13] Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering, 9*(3), 90–95. https://doi.org/10.1109/mcse.2007.55

[14] IJspeert, L. W., Tkachenko, A., Johnston, C., Garcia, S., De Ridder, J., Van Reeth, T., & Aerts, C. (2021). An all-sky sample of intermediate- to high-mass OBA-type eclipsing binaries observed by TESS. *Astronomy & Astrophysics*, 652, A120. https://doi.org/10.1051/0004-6361/202141489

[15] Kostov, V. B., Thompson, S. E., Quintana, E. V., Coughlin, J. L., Mullally, F., Barclay, T., Colón, K. D., Schlieder, J. E., Geert Barentsen, & Burke, C. J. (2019). Discovery and Vetting of Exoplanets. I. Benchmarking K2 Vetting Tools. *The Astronomical Journal, 157*(3), 124–124. https://doi.org/10.3847/1538-3881/ab0110

[16] Kovács, G., Zucker, S., & Mazeh, T. (2002). A box-fitting algorithm in the search for periodic transits. *Astronomy & Astrophysics*, 391(1), 369–377. https://doi.org/10.1051/0004-6361:20020802

[17] Kovács, G., Zucker, S., & Mazeh, T. (2016). BLS: Box-fitting Least Squares. Software, ascl 1607.008

[18] Kruse, E., Agol, E., Luger, R., & Foreman-Mackey, D. (2019). Detection of Hundreds of New Planet Candidates and Eclipsing Binaries in K2 Campaigns 0–8. *The Astrophysical Journal Supplement Series*, 244(1), 11. https://doi.org/10.3847/1538-4365/ab346b

[19] Mayor, M., & Queloz, D. (1995). A Jupiter-mass companion to a solar-type star. *Nature, 378*(6555), 355–359. https://doi.org/10.1038/378355a0

[20] Ofman, L., Averbuch, A., Shliselberg, A., Benaun, I., Segev, D., & Rissman, A. (2021). Automated identification of transiting exoplanet candidates in NASA Transiting Exoplanets Survey Satellite (TESS) data with machine learning methods. *New Astronomy*, 101693. https://doi.org/10.1016/j.newast.2021.101693

[21] Oliphant, T. E. (2007). Python for Scientific Computing. *Computing in Science & Engineering, 9*(3), 10–20. https://doi.org/10.1109/mcse.2007.58

[22]Olmschenk, G., Ishitani Silva, S., Rau, G., Barry, R. K., Kruse, E., Cacciapuoti, L., Kostov, V., Powell, B. P., Wyrwas, E., Schnittman, J. D., & Barclay, T. (2021). Identifying planetary transit candidates in TESS full-frame image light curves via convolutional neural networks. *The Astronomical Journal*, 161(6), 273. https://doi.org/10.3847/1538-3881/abf4c6

[23] Osborn, H. P., Ansdell, M., Ioannou, Y., Sasdelli, M., Angerhausen, D., Caldwell, D., Jenkins, J. M., Räissi, C., & Smith, J. C. (2020). Rapid classification of TESS planet candidates with convolutional neural networks. *Astronomy & Astrophysics*, 633, A53. https://doi.org/10.1051/0004-6361/201935345

[24] Perez, F., & Granger, B. E. (2007). IPython: A System for Interactive Scientific Computing. *Computing in Science & Engineering, 9*(3), 21–29. https://doi.org/10.1109/mcse.2007.53

[25] Powell, B. P., Kruse, E., Montet, B. T., Feinstein, A. D., Lewis, H. M., Foreman-Mackey, D., Barclay, T., Quintana, E. V., Colón, K. D., Kostov, V. B., Boyd, P., Smale, A. P., Mullally, S. E., Schlieder, J. E., Schnittman, J. D., Carroll, M. L., Carriere, L. E., Salmon, E. M., Strong, S. L., & Acks, N. D. (2022). The NASA GSFC TESS Full Frame Image Light Curve Data Set. *Research Notes of the AAS*, 6(6), 111. https://doi.org/10.3847/2515-5172/ac74c4

[26] Prsa, A., Kochoska, A., Conroy, K. E., Eisner, N. L., Hey, D. R., IJspeert, L. W., Kruse, E., Fleming, S. W., Johnston, C., Kristiansen, M. H., LaCourse, D., Mortensen, D., Pepper, J., Stassun, K. G., Torres, G., Abdul-Masih, M., Chakraborty, J., Gagliano, R., Guo, Z., & Hambleton, K. (2022). TESS Eclipsing Binary Stars. I. Short-cadence Observations of 4584 Eclipsing Binaries in Sectors 1–26. *Astrophysical Journal Supplement Series, 258*(1), 16–16. https://doi.org/10.3847/1538-4365/ac324a

[27] Ricker, G. R., Winn, J. N., Vanderspek, R., Latham, D. W., Bakos, Gáspár. Á., Bean, J. L., Berta-Thompson, Z. K., Brown, T. M., Buchhave, L., Butler, N. R., Butler, R. P., Chaplin, W. J., Charbonneau, D., Christensen-Dalsgaard, J., Clampin, M., Deming, D., Doty, J., De Lee, N., Dressing, C., & Dunham, E. W. (2014). Transiting Exoplanet Survey Satellite (TESS). *NASA ADS*, 9143, 914320. https://doi.org/10.1117/12.2063489

[28] Schmidt, B. (2011). Optical transient surveys. *Proceedings of the International Astronomical Union*, 7(S285), 9–10. https://doi.org/10.1017/s1743921312000129

APPENDIX

The following is a list of the 66 new confident EB candidates found using the pipeline. "Per" in the following table is short for period, all of which are measured in days. TICs with ECLIPSR periods of -1.0 failed its period-checking component. However, visual examination of their light curves resulted in their approval for the final list. Note that due to the naive nature of BLS, it is necessary to compare both half, normal, and double periods to check with the ECLIPSR period; if at least one matches within 5%, it can be concluded that the candidate is likely an EB.

Table 1.

TIC	ECLIPSR Score	BLS Max Power	ECLIPSR Per	BLS Per	BLS Half-Per	BLS Double-Per
142017247	0.70058	13681.907	3.77246	1.88869	0.944346	3.77739
142018009	0.74375	7363.6759	5.469861	2.736749	1.368375	5.473498
142087248	0.73804	6426.7324	3.206853	1.602473	0.8012367	3.204947
142106257	0.56804	8248.6442	-1.0	2.733216	1.366608	5.466431
142109975	0.41142	9653.9965	0.4671718	0.934629	0.4673145	1.869258
150427902	0.62584	9340.0726	0.5830927	0.5830389	0.2915194	1.166078
150428527	0.63894	22989.483	-1.0	17.21555	8.607774	34.4311
150430915	0.58194	16090.666	0.8496081	0.8498233	0.4249117	1.699647
150431387	0.5101	19722.816	0.7325885	0.7314488	0.3657244	1.462898
150437909	0.38911	217625.5	2.715104	1.35689	0.6784452	2.713781
150441162	0.56088	3629.7138	1.859403	0.9293286	0.4646643	1.858657
150442254	0.55518	141672.76	5.985096	1.489399	0.7446996	2.978799
150443182	0.86263	92283.073	2.599161	2.595406	1.297703	5.190813
150513899	0.67452	36595.109	3.767928	1.885159	0.9425795	3.770318
167246108	0.90129	40164.122	4.688481	2.344523	1.172261	4.689046
167250056	0.47302	67813.404	0.8584951	0.8586572	0.4293286	1.717314
167250519	0.55152	20382.502	1.249065	1.249117	0.6245583	2.498233
167251459	0.95775	566711.81	6.549728	3.273852	1.636926	6.547703
167308162	0.83735	5456.09	9.286412	9.30742	4.65371	18.61484
167308280	0.77423	48414.495	1.099269	0.54947	0.274735	1.09894
167308872	0.74702	36845.166	1.202029	1.199647	0.5998233	2.399293
167309182	0.70976	35595.374	3.993886	3.991166	1.995583	7.982332
167337786	0.84128	19726.025	8.432229	8.416961	4.208481	16.83392
167339237	0.76789	35216.537	4.369485	2.183746	1.091873	4.367491
167366499	0.70699	8735.5248	3.023812	3.021201	1.510601	6.042403
167421087	0.39361	7608.6522	0.557517	0.5583039	0.2791519	1.116608
260541900	0.68296	9532.2242	2.960863	1.480565	0.7402827	2.961131
260606384	0.78845	5391.5507	11.03146	11.01943	5.509717	22.03887
260608080	0.64695	7168.8303	3.424954	1.710247	0.8551237	3.420495
260608425	0.67785	262129.23	0.7632962	0.7632509	0.3816254	1.526502
260608573	0.70088	346398.22	1.388171	1.383392	0.6916961	2.766784
260635551	0.69275	34574.912	1.092206	0.5459364	0.2729682	1.091873
260636186	0.45719	3741.378	1.39404	1.388693	0.6943463	2.777385
260637574	0.67584	6431.9297	1.839197	0.9204947	0.4602473	1.840989
260640388	0.4711	5865.1884	0.4808455	0.9611307	0.4805654	1.922261

TIC	ECLIPSR Score	BLS Max Power	ECLIPSR Per	BLS Per	BLS Half-Per	BLS Double-Per
260653901	0.63321	16519.756	-1.0	12.63074	6.315371	25.26148
260654689	0.44783	7610.2046	1.281164	1.282686	0.6413428	2.565371
260658047	0.65166	6199.5947	0.9786507	0.9787986	0.4893993	1.957597
278683642	0.81095	28369.501	4.24603	4.238516	2.119258	8.477032
278683864	0.43504	5208.425	-1.0	18.85866	9.429329	37.71731
278683875	0.93839	520786.74	10.28662	18.79329	9.396643	37.58657
278684227	0.64573	17565.285	1.24816	0.6236749	0.3118375	1.24735
278726860	0.8238	4703.1256	9.213695	9.204947	4.602473	18.40989
278730069	0.72438	22200.132	2.486045	1.245583	0.6227915	2.491166
278776988	0.66254	51694.249	1.826993	0.9134276	0.4567138	1.826855
293220838	0.74858	37689.834	1.356033	1.355124	0.6775618	2.710247
293224945	0.82025	25822.334	3.570114	3.570671	1.785336	7.141343
293225157	0.8375	127256.87	5.573882	2.784452	1.392226	5.568905
293268007	0.77315	22873.484	4.061489	2.030035	1.015018	4.060071
293268574	0.75898	45415.106	5.115924	10.14488	5.072438	20.28975
293271277	0.8907	55029.873	4.567738	4.574205	2.287102	9.14841
293273040	0.7717	187060.5	4.960707	4.968198	2.484099	9.936396
293347945	0.48883	135326.7	-1.0	19.64664	9.823322	39.29329
348839218	0.79797	20275.583	2.970327	1.485866	0.7429329	2.971731
348839473	0.68513	5451.424	0.7573335	0.7579505	0.3789753	1.515901
348839494	0.8475	94841.57	3.174446	1.588339	0.7941696	3.176678
348841791	0.54564	23073.018	1.198034	1.19788	0.5989399	2.39576
348843066	0.63717	46696.295	-1.0	19.31802	9.659011	38.63604
372852412	0.73425	10213.969	2.001551	2.0	1.0	4.0
375033166	0.44292	4933.1203	0.4780668	0.9558304	0.4779152	1.911661
375035513	0.93495	864058.88	4.117609	2.056537	1.028269	4.113074
375058760	0.94502	2325567.2	7.137274	7.132509	3.566254	14.26502
514257394	0.64561	7174.3868	3.4246	1.710247	0.8551237	3.420495
737153831	0.93495	864059.02	4.117609	2.056537	1.028269	4.113074
737487643	0.75251	8104.8528	0.7573412	0.7579505	0.3789753	1.515901
737496372	0.65166	6199.5928	0.9786507	0.9787986	0.4893993	1.957597