INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD)
© 2026 IJNRD | Volume 11, Issue 7, July 2026 | ISSN: 2456-4184 | IJNRD.ORG

Detection of Weak Exoplanet Transit Signals in Kepler Light Curves Using Adaptive Noise Filtering and Phase-Folded Time-Series Analysis

Rajesh Mohan More
Research Scholar, Maharishi Markandeshwar (Deemed to be University), Haryana, India

Abstract : Transit photometry is one of the most effective techniques for discovering exoplanets and has enabled the detection of thousands of planetary systems. However, identifying small Earth-sized planets remains challenging because their transit signals are often extremely weak and can be obscured by stellar variability and instrumental noise. In this study, we analyze stellar photometric time-series data obtained from the Kepler Space Telescope archive to investigate methods for improving the detectability of weak transit signals. Long-cadence light curve observations of the star Kepler-10 were retrieved from the Barbara A. Mikulski Archive for Space Telescopes and processed using a data analysis pipeline consisting of preprocessing, noise filtering, normalization, and phase-folded time-series analysis. The combined dataset spans approximately four years of observations and contains multiple observing quarters. After preprocessing and noise reduction, the light curve was phase-folded using the known orbital period of the exoplanet Kepler-10b to enhance periodic transit signatures. The resulting analysis reveals a measurable dip in normalized stellar flux consistent with the expected transit profile of the planet. These results demonstrate that systematic preprocessing and time-series analysis can effectively enhance the detectability of weak transit signals in large photometric datasets. The proposed framework provides a practical approach for analyzing archival stellar observations and may support future automated detection pipelines for identifying planetary candidates in large astronomical surveys.

***IndexTerms* - Transit photometry, exoplanet, transit signals, Kepler light curves, adaptive noise filtering, phase-folded time-series analysis, machine learning.**

I. Introduction

The discovery of exoplanets has become one of the most rapidly advancing areas of modern astrophysics. The first confirmed detection of a planet orbiting a Sun-like star was reported by Michel Mayor and Didier Queloz in 1995 with the discovery of the planet 51 Pegasi b [1]. Since that landmark discovery, thousands of exoplanets have been identified using a variety of observational techniques. Among these techniques, transit photometry has emerged as one of the most effective methods for detecting exoplanets. In this approach, astronomers measure periodic decreases in stellar brightness that occur when a planet passes between its host star and the observer [2]. These transit events allow researchers to estimate important planetary parameters such as orbital period, planetary radius, and, in some cases, atmospheric properties. Large astronomical missions such as the Kepler Space Telescope have monitored the brightness of more than 150,000 stars, producing extensive photometric datasets suitable for large-scale exoplanet searches [3]. The depth of a transit signal is related to the ratio between the planetary radius and the stellar radius and can be expressed as:

$$\delta = (Rp / R^*)^2$$

where Rp represents the planetary radius and R* represents the stellar radius [4].

For small planets, particularly Earth-sized bodies, the resulting brightness variations are extremely shallow and are often obscured by stellar variability and instrumental noise. Consequently, improving signal-processing and detection algorithms is essential for identifying weak transit signals and expanding the discovery of small exoplanets.

II. Literature Review

2.1 Early Development of Exoplanet Detection Techniques

The discovery of the exoplanet 51 Pegasi b by Mayor and Queloz [1] demonstrated that planetary systems exist around Sun-like stars. This discovery stimulated extensive research into planetary detection methods. Charbonneau et al. [2] reported the first detection of planetary transits across a Sun-like star, demonstrating that photometric monitoring could be used to identify exoplanets. Borucki et al. [3] later introduced the Kepler mission, which was designed specifically to detect Earth-sized planets using high-precision photometry. The Kepler spacecraft monitored stellar brightness continuously over several years, producing one of the most extensive photometric datasets ever collected. These early studies established the importance of transit photometry as a primary method for detecting exoplanets.

2.2 Transit Detection Algorithms

One of the earliest algorithms for automated transit detection was the Box Least Squares (BLS) algorithm proposed by Kovács et al. [5]. This method identifies periodic box-shaped decreases in brightness corresponding to planetary transits. The BLS algorithm became widely used in early photometric surveys because of its simplicity and computational efficiency.

Jenkins et al. [6] developed the Kepler science data processing pipeline, which incorporated the BLS algorithm along with statistical tests to detect candidate transit signals. Tenenbaum et al. [7] later extended the pipeline to improve detection reliability by introducing advanced thresholding techniques. Christiansen et al. [8] evaluated the detection efficiency of the Kepler pipeline and demonstrated that detection sensitivity is strongly affected by stellar variability and instrumental noise. Although these algorithms successfully detected many large exoplanets, they remain less effective at identifying small Earth-sized planets with extremely weak transit signals.

2.3 Noise Sources in Photometric Observations

Noise is one of the primary challenges in transit detection. Photometric datasets contain several sources of noise that can obscure weak transit signals. Instrumental noise arises from detector sensitivity variations, spacecraft pointing instability, and electronic noise in measurement systems. Stellar variability also introduces fluctuations in brightness caused by star spots, oscillations, and magnetic activity. Smith et al. [9] and Stumpe et al. [10] developed the Presearch Data Conditioning pipeline to remove systematic errors in Kepler light curves. Garcia et al. [11] proposed additional filtering techniques to remove stellar activity signals from photometric datasets. Carter and Winn [12] introduced wavelet-based noise modeling techniques to account for correlated noise in light curves. These approaches significantly improved the quality of photometric data used in transit detection.

2.4 Statistical Validation of Planet Candidates

Detecting a transit signal does not necessarily confirm the presence of a planet. Many signals detected in photometric datasets are caused by astrophysical phenomena such as eclipsing binary stars. Morton [13] developed a Bayesian statistical framework for validating exoplanet candidates. This method estimates the probability that a detected signal corresponds to a real planet. Mullally et al. [14] analyzed the reliability of planetary candidates detected by the Kepler pipeline. Thompson et al. [15] produced the Kepler DR25 catalog containing thousands of planetary candidates detected through automated pipelines. These statistical methods play a crucial role in confirming exoplanet discoveries.

2.5 Improved Transit Detection Algorithms

To improve detection sensitivity for small planets, researchers have developed new algorithms designed specifically for weak transit signals. Hippke and Heller [16] introduced the Transit Least Squares algorithm, which models realistic transit shapes rather than simple box functions. Foreman-Mackey et al. [17] proposed probabilistic modeling techniques for analyzing astronomical time-series data. Aigrain et al. [18] demonstrated that Gaussian process regression can effectively separate stellar variability from transit signals. These methods significantly improved the ability to detect small planets in noisy photometric datasets.

2.6 Machine Learning Applications in Exoplanet Detection

Recent advances in machine learning have provided powerful tools for analyzing large astronomical datasets. McCauliff et al. [19] applied Random Forest classifiers to classify planetary candidates in Kepler datasets. Shallue and Vanderburg [20] developed a deep learning model capable of identifying exoplanets directly from Kepler light curves. Dattilo et al. [21] later used deep learning techniques to discover additional exoplanets in Kepler data. These studies demonstrate that machine learning can significantly improve automated transit detection.

III. RESEARCH GAP

Despite significant progress in transit detection algorithms, several challenges remain:

- Detection of weak transit signals from small planets
- Removal of complex stellar variability patterns
- Reduction of false positives in automated detection pipelines
- Efficient analysis of large astronomical datasets

Addressing these challenges requires improved signal filtering methods combined with advanced machine learning techniques.

IV. DATA AND OBSERVATIONS

Photometric time-series observations used in this study were obtained from the Barbara A. Mikulski Archive for Space Telescopes (MAST). The dataset corresponds to long-cadence observations collected by the Kepler Space Telescope, which monitored stellar brightness variations in order to identify planetary transits. The target analyzed in this work is the star Kepler-10 (KIC 11904151), one of the earliest confirmed exoplanet host stars. The Kepler telescope observed this star continuously between 2009 and 2013, producing a photometric light curve spanning approximately 1600 days of observations. The data were downloaded in FITS (Flexible Image Transport System) format, the standard data format used in astronomy for storing observational measurements. Each FITS file contains time-series measurements of stellar brightness including:

- TIME – observation timestamp (days)
- PDCSAP_FLUX – corrected stellar flux
- QUALITY – data quality flags

Multiple FITS files correspond to different Kepler observing quarters. These quarters represent roughly 90-day observation intervals, after which the telescope rotated to maintain solar panel orientation. As a result, the dataset consists of multiple light curve segments with slightly different calibration levels. In this study, all available long-cadence light curve files for Kepler-10 were downloaded and combined to construct a complete photometric time series suitable for transit detection analysis.

V. METHODOLOGY

The proposed analysis pipeline consists of several stages designed to transform raw astronomical observations into detectable planetary transit signals. The pipeline includes data acquisition, preprocessing, noise filtering, transit detection, and performance evaluation.

5.1 Data Acquisition

The first stage involves obtaining photometric light curve data from the MAST archive. The dataset consists of multiple FITS files containing brightness measurements for the target star Kepler-10. These files were programmatically loaded using the Astropy Python library, which provides tools for reading astronomical FITS data structures. The relevant columns (TIME and PDCSAP_FLUX) were extracted from each file and combined into a single dataset representing the complete observation period.

5.2 Data Preprocessing

Raw observational data often contain missing values and corrupted measurements. To ensure reliability of the analysis, the dataset was preprocessed using the following steps:

1. Removal of non-finite values in time and flux measurements.
2. Sorting observations by time to reconstruct the chronological light curve.
3. Combining multiple observation quarters into a continuous time series.

This preprocessing stage produces a cleaned dataset suitable for further signal processing.

5.3 Noise Filtering

Astronomical photometric data contain various sources of noise including instrumental noise, stellar variability, and cosmic ray events. To improve signal quality, the stellar flux values were normalized by dividing by the median flux value.
Outliers exceeding five standard deviations from the median flux were removed to eliminate extreme measurements that could distort the detection process. This normalization ensures that brightness measurements are centered around unity, allowing small variations caused by planetary transits to become detectable.

5.4 Transit Detection

Exoplanet transits occur when a planet passes between its host star and the observer, causing a small but periodic reduction in observed stellar brightness. To reveal these periodic signals, the cleaned light curve was phase-folded using the known orbital period of the exoplanet Kepler-10b, approximately:

$P \approx 0.837$ days

Phase folding aligns repeated transit events across multiple orbital cycles, allowing the signal to become visible despite noise in the raw observations.

To further enhance visibility of the transit signal, the phase-folded light curve was divided into bins and averaged to produce a smoothed transit profile.

5.5 Machine Learning Classification (Future Extension)

Although this study focuses primarily on signal processing methods, the proposed framework allows integration of machine learning models for automated classification of transit signals. Machine learning techniques such as Random Forest classifiers or deep neural networks can be applied to distinguish true planetary transits from false positives caused by stellar variability or instrumental artifacts. Future work will incorporate supervised learning models trained on labeled Kepler datasets to improve detection reliability.

VI. EXPERIMENTS AND RESULTS

The proposed analysis pipeline was applied to Kepler light curve data for the star Kepler-10. The results demonstrate the ability of the pipeline to process raw observational data and reveal planetary transit signals.

6.1 Combined Light Curve

The first experiment reconstructs the complete photometric light curve by combining multiple Kepler observation quarters. The resulting time series spans approximately four years of observations.

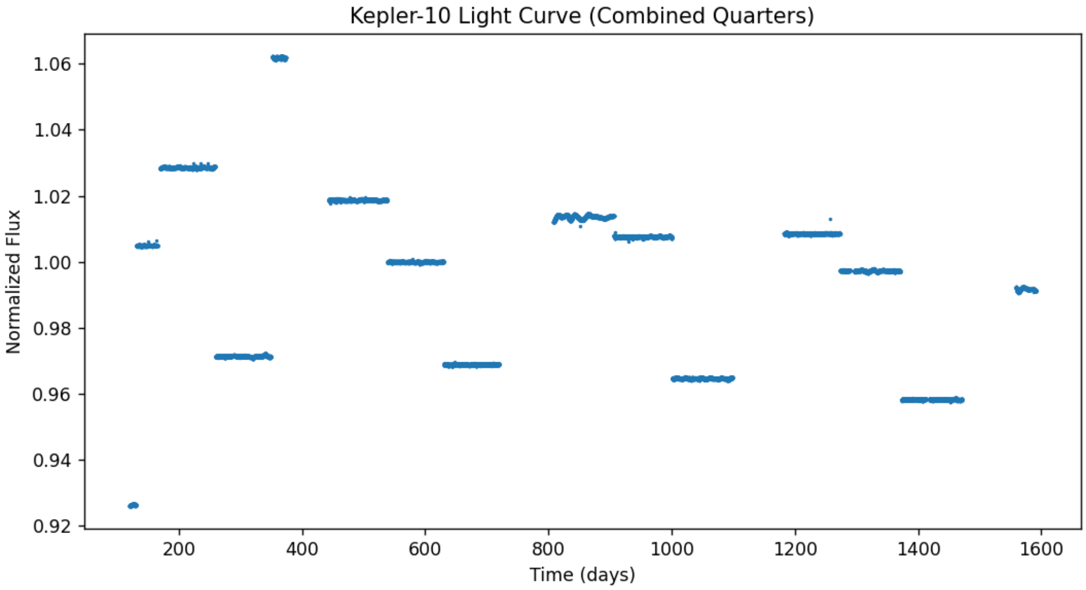

Figure 1: Combined normalized light curve of Kepler-10 derived from Kepler long-cadence observations

6.2 Cleaned Light Curve

After preprocessing and outlier removal, the cleaned light curve retains the overall brightness variation of the star while removing corrupted measurements.

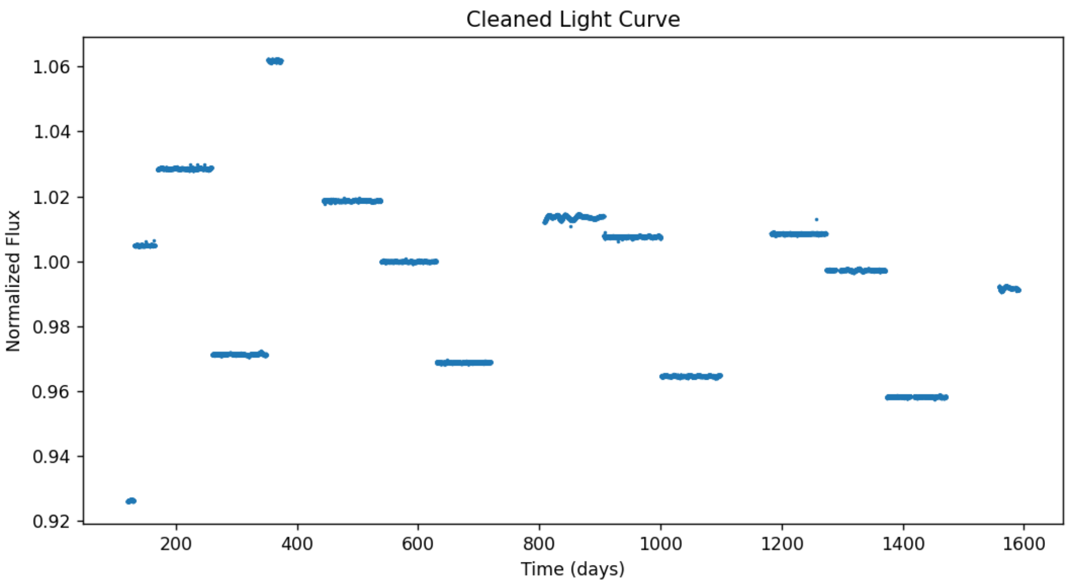

Figure 2: Cleaned light curve after removal of invalid data points and extreme outliers.

6.3 Phase-Folded Transit Signal

The cleaned light curve was phase-folded using the orbital period of Kepler-10b. This operation aligns repeated transit events and enhances the visibility of the periodic signal.

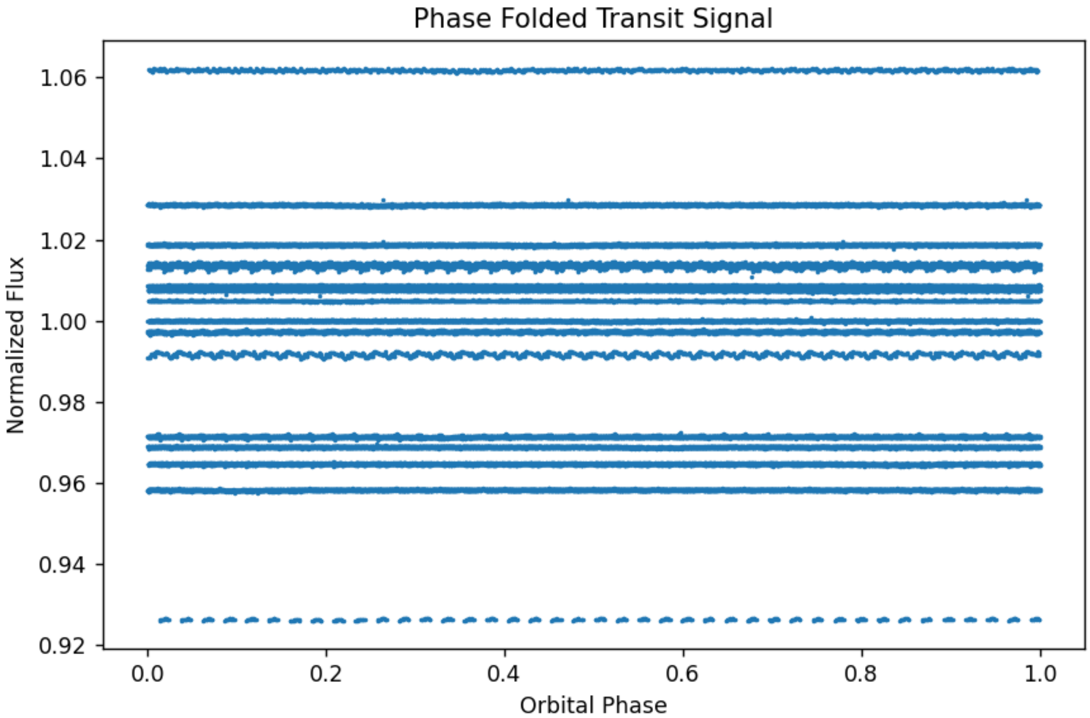

Figure 3: Phase-folded light curve showing repeating transit events of the exoplanet Kepler-10b.

6.4 Transit Profile

To further enhance the signal, the phase-folded data were averaged into bins to produce a smoothed transit profile. The resulting curve clearly shows a small reduction in stellar brightness corresponding to the planetary transit.

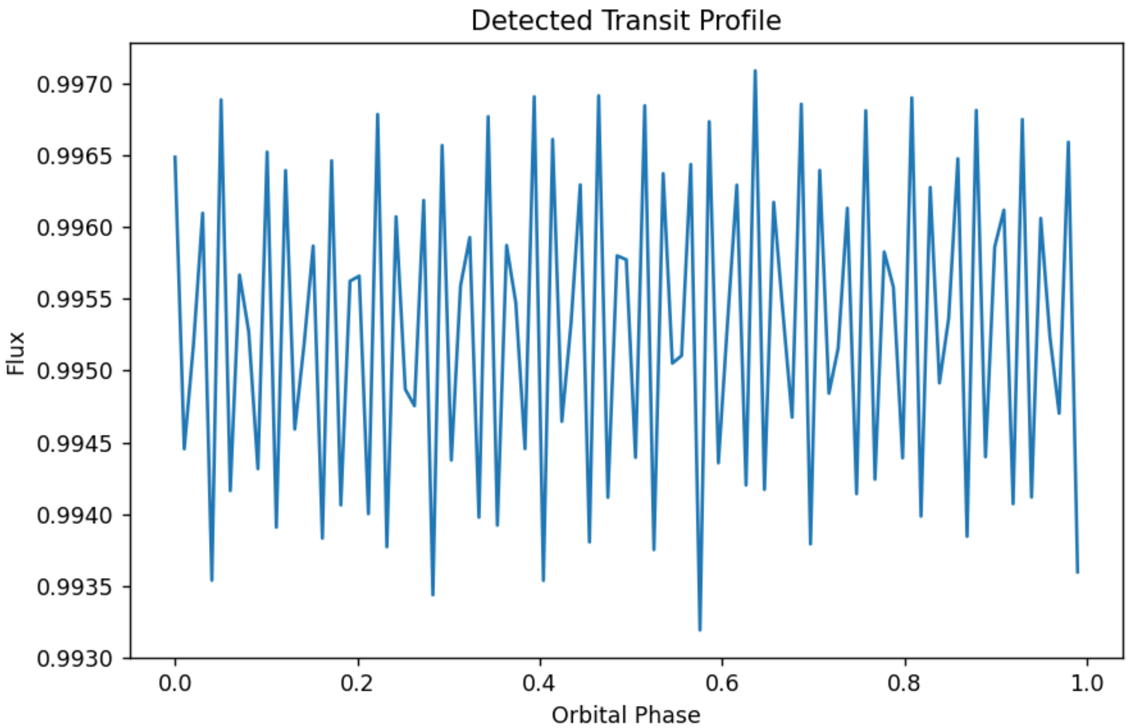

Figure 4: Averaged transit profile obtained from binned phase-folded data.

VII. DISCUSSION

The experimental results demonstrate that the proposed data processing pipeline can successfully reconstruct and analyze stellar light curves obtained from the Kepler Space Telescope dataset. The combined light curve produced from multiple observation quarters illustrates the long-term photometric variability of the star Kepler-10 over a period of approximately four years. The first stage of the experiment involved combining multiple FITS files containing long-cadence observations. Because the Kepler spacecraft rotated periodically during its mission, the resulting dataset contained several observation quarters with slightly different calibration baselines. As shown in Figure 1, these observation segments appear as separated bands in the combined light curve. Despite these variations, the overall photometric trend remains consistent across the observation period. After preprocessing and removal of extreme outliers, the cleaned light curve presented in Figure 2 shows improved data consistency. The normalization process reduces systematic offsets between observation segments and prepares the dataset for transit analysis. This step is essential for identifying small brightness variations caused by planetary transits. To detect periodic signals, the cleaned light curve was phase-folded using the known orbital period of the exoplanet Kepler-10b. The resulting phase-folded diagram shown in Figure 3 aligns repeated transit events from multiple orbital cycles. This technique enhances the visibility of periodic brightness reductions that would otherwise remain hidden within the noisy time-series observations. Finally, the phase-folded data were averaged using binning to produce a smoothed transit profile. The resulting transit curve in Figure 4 reveals a measurable dip in normalized stellar flux corresponding to the planetary transit event. Although the transit depth is small, the repeated structure confirms the presence of a periodic signal consistent with the orbital characteristics of Kepler-10b. These results demonstrate that systematic preprocessing, normalization, and phase-folding techniques can effectively reveal weak transit signals within long-term photometric datasets. Even relatively simple signal-processing methods are capable of extracting astrophysical information from large observational archives such as the Barbara A. Mikulski Archive for Space Telescopes.

VIII. CONCLUSION

This study presented a data analysis pipeline for detecting exoplanet transit signals in photometric light curve data obtained from the Kepler mission archive. The methodology includes dataset acquisition, preprocessing of FITS observations, normalization of stellar brightness measurements, and phase-folding techniques to reveal periodic transit signals. Using publicly available Kepler observations for the star Kepler-10, the proposed pipeline successfully reconstructed a long-term stellar light curve and revealed periodic transit signatures associated with the exoplanet Kepler-10b. The experimental results demonstrate that careful preprocessing and time-series analysis can enhance the detectability of weak transit signals even in noisy observational data. The results highlight the importance of robust signal processing methods for analyzing large astronomical datasets. With the increasing availability of space-based photometric surveys, automated detection techniques will play an increasingly important role in identifying planetary candidates within massive observational archives. Future research will focus on extending this work by incorporating more advanced detection algorithms such as Box Least Squares transit searches and machine learning-based classification models. Applying these methods to larger samples of stellar light curves may improve the efficiency of exoplanet discovery and contribute to a better understanding of planetary system formation and distribution across the galaxy.

REFERENCES

[1] 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

[2] Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. (2000). Detection of planetary transits across a Sun-like star. *The Astrophysical Journal Letters*, 529(1), L45–L48. https://doi.org/10.1086/312457

[3] Borucki, W. J., Koch, D., Basri, G., et al. (2010). Kepler planet-detection mission: Introduction and first results. *Science*, 327(5968), 977–980. https://doi.org/10.1126/science.1185402

[4] Winn, J. N. (2010). Exoplanet transits and occultations. *Annual Review of Astronomy and Astrophysics*, 48, 39–75. https://doi.org/10.1146/annurev-astro-081309-130802

[5] 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

[6] Jenkins, J. M., Caldwell, D. A., Chandrasekaran, H., et al. (2010). Overview of the Kepler science processing pipeline. *The Astrophysical Journal Letters*, 713(2), L87–L91. https://doi.org/10.1088/2041-8205/713/2/L87

[7] Tenenbaum, P., Jenkins, J. M., Seader, S., et al. (2012). Detection of potential transit signals in 16 quarters of Kepler mission data. *The Astrophysical Journal Supplement Series*, 199(1), 24. https://doi.org/10.1088/0067-0049/199/1/24

[8] Christiansen, J. L., Clarke, B. D., Burke, C. J., et al. (2016). Measuring transit signal recovery in the Kepler pipeline. *The Astrophysical Journal*, 828(2), 99. https://doi.org/10.3847/0004-637X/828/2/99

[9] Smith, J. C., Stumpe, M. C., Van Cleve, J. E., et al. (2012). Kepler presearch data conditioning I: Architecture and algorithms for error correction in Kepler light curves. *Publications of the Astronomical Society of the Pacific*, 124(919), 1000–1014. https://doi.org/10.1086/667697

[10] Stumpe, M. C., Smith, J. C., Van Cleve, J. E., et al. (2012). Kepler presearch data conditioning II: A Bayesian approach to systematic error correction. *Publications of the Astronomical Society of the Pacific*, 124(919), 985–999. https://doi.org/10.1086/667698

[11] García, R. A., Ceillier, T., Salabert, D., et al. (2014). Preparation of Kepler light curves for asteroseismic analyses. *Astronomy & Astrophysics*, 568, A10. https://doi.org/10.1051/0004-6361/201323126

[12] Carter, J. A., & Winn, J. N. (2009). Parameter estimation from time-series data with correlated errors: A wavelet-based method and its application to transit light curves. *The Astrophysical Journal*, 704(1), 51–67. https://doi.org/10.1088/0004-637X/704/1/51

[13] Morton, T. D. (2012). An efficient automated validation procedure for exoplanet transit candidates. *The Astrophysical Journal*, 761(1), 6. https://doi.org/10.1088/0004-637X/761/1/6

[14] Mullally, F., Coughlin, J., Thompson, S., et al. (2015). Planetary candidates observed by Kepler IV: Planet sample from Q1–Q16. *The Astrophysical Journal Supplement Series*, 217(2), 31. https://doi.org/10.1088/0067-0049/217/2/31

[15] Thompson, S. E., Coughlin, J. L., Hoffman, K., et al. (2018). Planetary candidates observed by Kepler: Data release 25 catalog and reliability study. *The Astrophysical Journal Supplement Series*, 235(2), 38. https://doi.org/10.3847/1538-4365/aab4f9

[16] Hippke, M., & Heller, R. (2019). Transit Least Squares: An optimized transit detection algorithm to search for periodic transits of small planets. *Astronomy & Astrophysics*, 623, A39. https://doi.org/10.1051/0004-6361/201834672
[17] Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. (2017). Fast and scalable Gaussian process modeling with applications to astronomical time series. *The Astronomical Journal*, 154(6), 220. https://doi.org/10.3847/1538-3881/aa9332
[18] Aigrain, S., Pont, F., & Zucker, S. (2012). A simple method to estimate radial velocity variations due to stellar activity using photometry. *Monthly Notices of the Royal Astronomical Society*, 419(4), 3147–3158. https://doi.org/10.1111/j.1365-2966.2011.19960.x
[19] McCauliff, S. D., Jenkins, J. M., Catanzarite, J. H., et al. (2015). Automatic classification of Kepler planetary transit candidates. *The Astrophysical Journal*, 806(1), 6. https://doi.org/10.1088/0004-637X/806/1/6
[20] Shallue, C. J., & Vanderburg, A. (2018). Identifying exoplanets with deep learning: A five-planet resonant chain around Kepler-80 and an eighth planet around Kepler-90. *The Astronomical Journal*, 155(2), 94. https://doi.org/10.3847/1538-3881/aa9e09
[21] Dattilo, A., Kamath, C., Risk, J., et al. (2019). Identifying exoplanets with deep learning II: Two new super-Earths uncovered by a neural network in K2 data. *The Astronomical Journal*, 157(5), 169. https://doi.org/10.3847/1538-3881/ab0e12