Introducing FLATW'RM

A machine-learning tool for finding flares in *Kepler* data

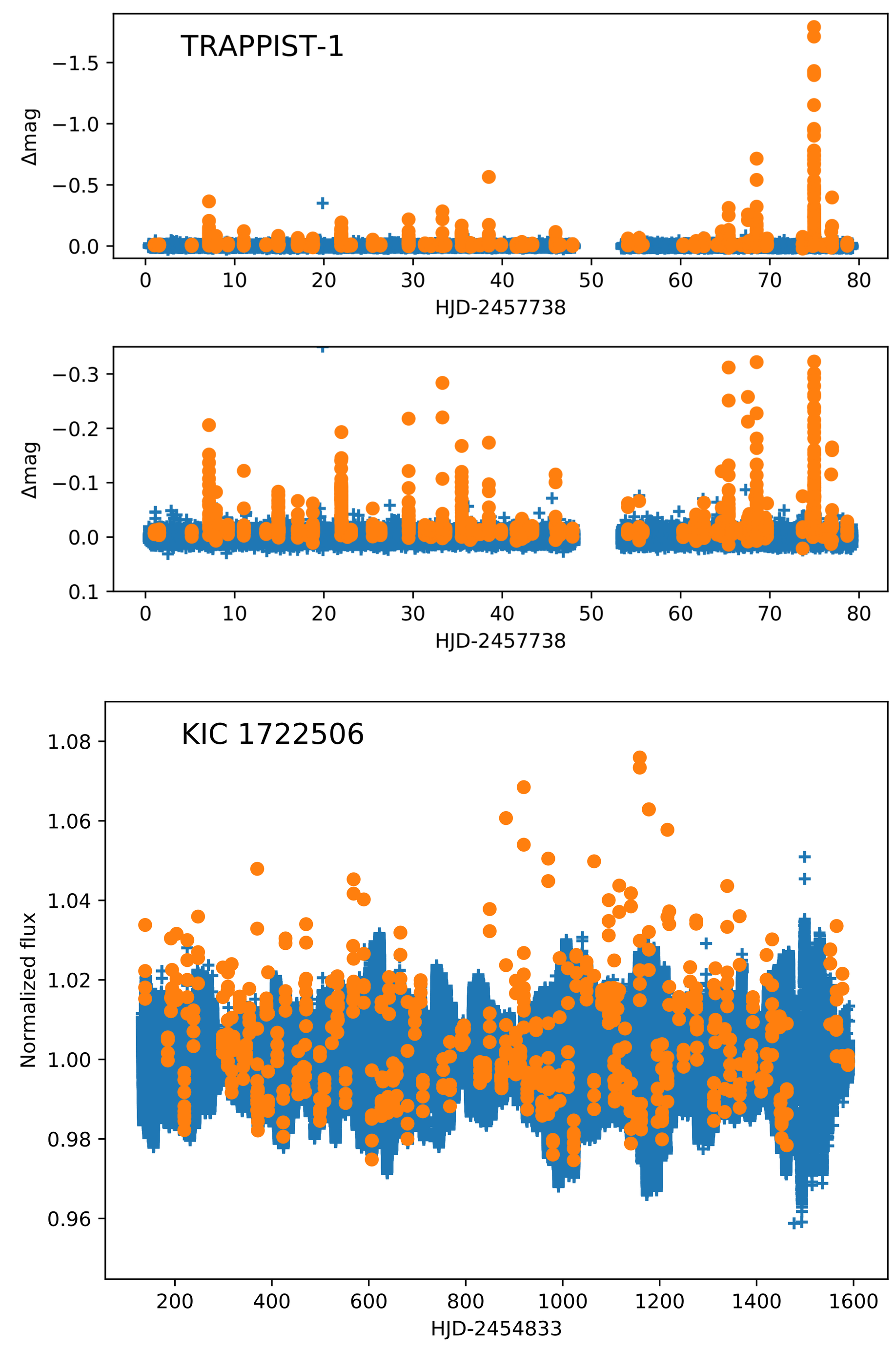
K. Vida[1], R. Roettenbacher[2,3]

vidakris@konkoly.hu

[1] Konkoly Observatory, MTA CSFK, Budapest, Hungary; [2] Department of Astronomy, Stockholm University, SE-106 91 Stockholm, Sweden; [3] Department of Astronomy, Yale University, PO Box 208101, New Haven, CT, USA

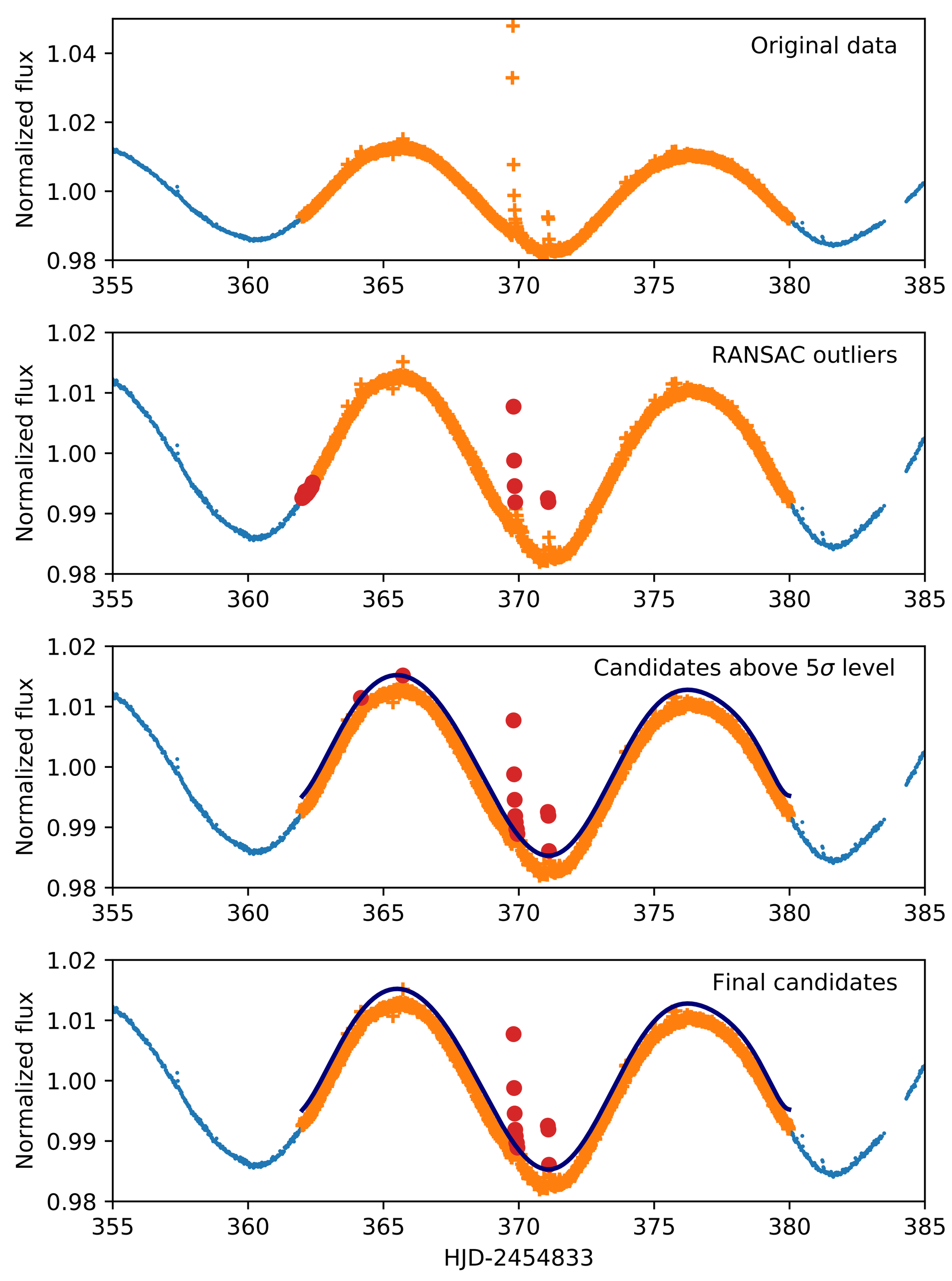
Archives of long photometric surveys, like the *Kepler* database, are a gold mine for studying flares. However, identifying them is a complex task; while in the case of single-target observations it can be easily done manually by visual inspection, this is nearly impossible for years-long time series for several thousand targets. Although there exist automated methods for this task, several problems are difficult (or impossible) to overcome with traditional fitting and analysis approaches.

We introduce the FLATW'RM code for identifying and analyzing flares based on machine-learning methods, which are intrinsically adept at handling such data sets. The RANSAC (RANdom SAmple Consensus) algorithm was used to model light curves, as it yields robust fits even in case of several outliers, like flares. The code was tested on the short-cadence *K2* observations of TRAPPIST-1 and on the long-cadence *Kepler* data of KIC 1722506.



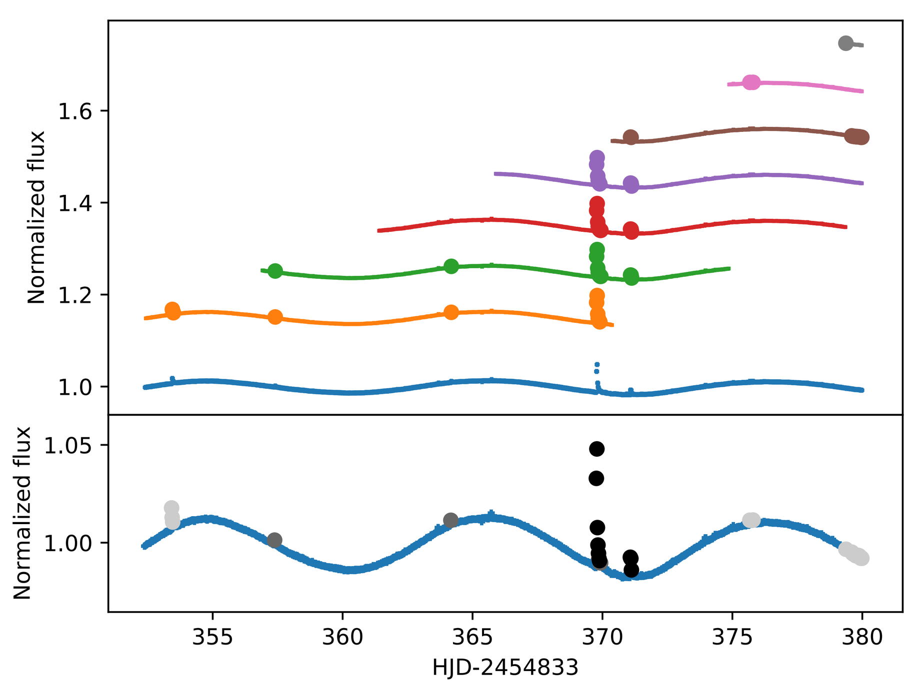
Top: Selected flare candidates in the TRAPPIST-1 short-cadence K2 data. The upper panel shows the total light curve, the middle plot is zoomed-in to show smaller events. Bottom: Similar analysis, but for long-cadence Kepler data of KIC 1722506.

Selecting flare candidates

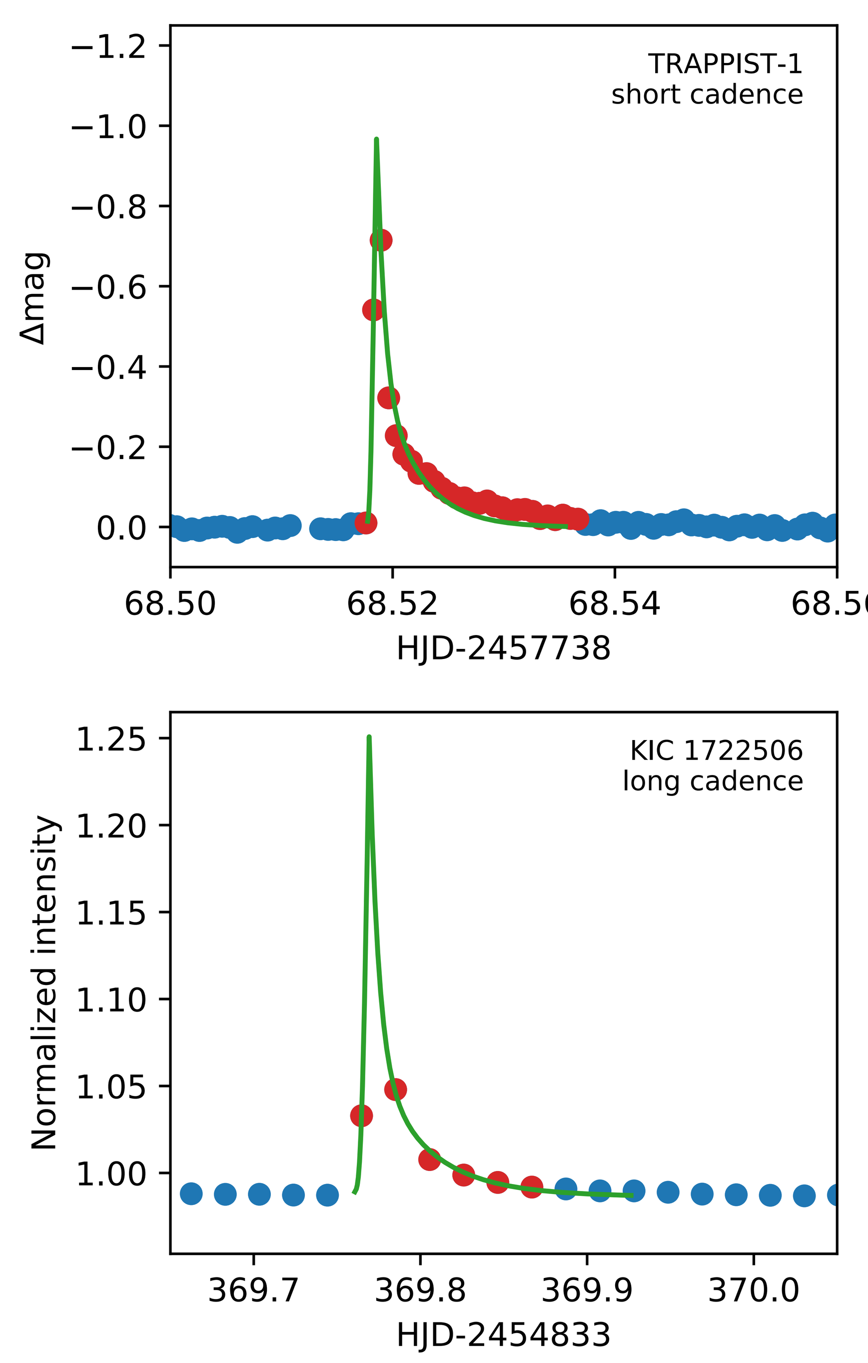


Candidate selection: the top plot shows the original light-curve section. The second plot shows the outlier points found by RANSAC marked with red dots. The third plot shows the 5σ detection level from the fit (continuous line) and the flare candidate points. In the bottom plot, the final flare candidates are shown, which have more than a given number (three, in this case) of consecutive data points. These points will get a vote for this light-curve section indicating that the feature likely is a flare.

Demonstration of the algorithm on a light-curve section of KIC 1722506. The light curve is divided into search windows (left), approximately on the order of the stellar rotation period. This window is shifted over the dataset, and a voting system is used to keep false positives to a minimum: only those flare candidate points are kept that were identified in several windows as a flare (bottom).



Voting algorithm: The top plot shows the original light curve (in blue, near a normalized flux of 1), and each light-curve segment in individual windows tested by FLATW'RM (each a different color), and the flare candidates are marked for the given run with circles. In the bottom plot, the candidates with one (light gray), two (medium gray), and at least three (black) votes are plotted. In this setup, the flares plotted in black are kept as final flare candidates.



Two samples from the recovered flares, from the short-cadence TRAPPIST-1 and the long-cadence KIC 1722506 data. Red points show the points selected for flares, the green line indicates the fitted analytical model from Davenport et al. (2014, ApJ, 797, 122)

Acknowledgements

The authors acknowledge the Hungarian National Research, Development and Innovation Office grants OTKA K-109276, OTKA K-113117. KV is supported by the Bolyai János Research Scholarship of the Hungarian Academy of Sciences. This work has used *K2* data from the proposal number GO12046. Funding for the *Kepler* and *K2* missions is provided by the NASA Science Mission directorate.

Get the code!

FLATW'RM

Flare detection with RANSAC method

Interested? Read the paper! arXiv:1806.00334

XXXth General Assembly of the International Astronomical Union
Vienna, 2018 August 20–31