Theoretical and Computational Methods

Identifying Stellar Flares From TESS Lightcurves

Teo Korhonen

March 27, 2020

Supervisor(s):

Examiner(s):

University of Helsinki
Faculty of Science

PL 64 (Gustaf Hällströmin katu 2a)
00014 Helsingin yliopisto

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta — Fakultet — Faculty	Koulutusohjelma — Utbildningsprogram — Degree programme
Faculty of Science	Theoretical and Computational Methods

Tekijä — Författare — Author

Teo Korhonen

Työn nimi — Arbetets titel — Title

Identifying Stellar Flares From TESS Lightcurves

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidantal — Number of pages
	March 27, 2020	13

Tiivistelmä — Referat — Abstract

Avainsanat — Nyckelord — Keywords

Säilytyspaikka — Förvaringsställe — Where deposited

Muita tietoja — Övriga uppgifter — Additional information

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta — Fakultet — Faculty	Koulutusohjelma — Utbildningsprogram — Degree programme
Faculty of Science	Theoretical and Computational Methods

Tekijä — Författare — Author

Teo Korhonen

Työn nimi — Arbetets titel — Title

Identifying Stellar Flares From TESS Lightcurves

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidantal — Number of pages
	March 27, 2020	13

Tiivistelmä — Referat — Abstract

Avainsanat — Nyckelord — Keywords

Säilytyspaikka — Förvaringsställe — Where deposited

Muita tietoja — Övriga uppgifter — Additional information

Contents

1. Introduction

This thesis explores the detection of stellar flares from the lightcurves of a selection of highly active stars.

1.1 Motivation

Stellar flares are powerful energy bursts originating from the surface of a star. They are commonly thought to be the result of the reorganization of the magnetic fields on the surface of the star. Most of our knowledge comes from solar surveys through probes and telescopes, which has given us a detailed picture of the magnetic activity occurring on the surface of the Sun.

Flares from other stars have been subject to less research. Analyzing these flares is much more challenging, since the only consequence of a stellar flare visible to current instruments is a momentary variation in flux. Detecting these flares has been previously It requires long enough lightcurves to contain significant flares in the first place. In addition, the lightcurve needs to be of sufficiently high resolution for the flare to be visible. Statistical noise, interference from nearby stars, and other types of anomalies add to the challenge of identifying a possible flare signal. Notwithstanding, the last decade has seen a surge in research into stellar flares.

Particular challenge arises from identifying flare events when the lightcurve is highly variable. Magnetically active stars can have short-term multiple periodicity due to sunspots and differential rotation. Often, this periodicity can take place in nearly the same timescale as the flares themselves. This means that traditional de-trending approaches may be ill-conditioned for stars of particularly high magnetic activity.

1.2 Overview of the literature

With a surge in exoplanet surveys, an unprecedented amount of detailed lightcurve data has been recorded on a large number of stars. Among those surveys is NASA's Transiting Exoplanet Survey Satellite (TESS) [Ricker, 2014]. TESS launched in 2018, and has since been collecting lightcurves from a large number of stars across wide areas of the sky in pursuit of exoplanet transits. NASA anticipates TESS to discover almost 15,000 exoplanets [Barclay et al., 2018]. In doing so, it produces a large set of detailed, high-cadence stellar lightcurves suitable for detection of flares. Similar flare research has been conducted on the data from previous exoplanet missions Kepler and K2, as well as dedicated lightcurve measurements.

One important finding from stellar flare research is the existence of *superflares*, which are stellar flares that are orders of magnitude more energetic than the largest flares ever detected on the Sun [Schaefer et al., 2000]. The possibility and the potential consequences of a superflare occurring on the Sun are a concern for the continued function of e.g. satellites and electric grids on Earth. This motivates much of the research on stellar flares, their causes, and their rates of occurrence. Earlier research associated superflares to interactions with twin stars and "hot Jupiter"-type planets in close orbit with the star, [Rubenstein and Schaefer, 2008] but newer studies have found superflares with no evidence of such orbiters. [Maehara, 2012] While it appears that superflares generally occur on stars with a considerably higher magnetic activity than the Sun, many results suggest that they might still be possible, if rare, on a solar-type star [Tu et al., 2020].

Other findings from extrasolar flares include the relationship between spectral class and flare activity, which has been confirmed across literature [Doorsselaere et al., 2017]. Similarly, there appears to be a separate group of highly active flare stars from all spectral classes.

Different methods for detecting the flares have been used across the research. Most earlier studies, such as Schaefer, have largely used manual inspection to detect flare candidates, and then excluded likely false positives based on various types of information about the observations. The manual approach has produced much of our knowledge about stellar flares, but it is not rigorous and largely impractical for larger datasets. Especially with complex and rapidly varying lightcurves, false positives and false negatives are very likely with human inspection.

Maehara's approach starts from the jumps between successive lightcurve datapoints, and then excludes likely false positives by removing the long-term trend from the data. The long-term trend was determined by fitting a polynomial based on datapoints before and after the suspected flare. The authors still had to manually inspect the results and remove some false signals, particularly near gaps in the dataset.

Some studies have attempted a totally systematic approach. Doorsselaere's paper found flares in Kepler's long-cadence lightcurve data by fitting a 3rd order polynomial and various periodic signals to longer lightcurve segments (sans statistical outliers), and then considering datapoints far from the fit as flare candidates. Davenport's catalogue of Kepler flares [Davenport, 2016], across both short- and long-cadence Kepler lightcurves, uses a similarly complicated detrending procedure. This approach largely found high-energy flares ($\approx 10^{35}$ erg, vs. $\approx 10^{30}$ for most solar flares).

Lastly, one study applied a robust machine learning approach to fit a model of the lightcurve and detected flare outliers. [Vida and Roettenbacher, 2018] This algorithm, nicknamed FLATW'RM, uses a rolling window of the lightcurve, and iteratively picks a random sample of datapoints from each window to find the most robust fit. The flare candidates are cross-validated across overlapping windows. The method was successful and produced similar results to hand-picked flare candidates. For the purposes of this work, adapting an approach similar to FLATW'RM is the most appealing; the high variability in active star lightcurves are likely ill-conditioned for the longer-window fits in the other papers. A shorter-term, robust fit is necessary to avoid issues caused by complex periodicity.

1.3 Scope and structure

The primary purpose of this thesis is to build a systematic approach to detect flares from the lightcurves of active stars, using various methods for time series analysis. These methods are subsequently used to detect flares from a selection of magnetically active stars from TESS lightcurves.

Chapter 2 covers the current theoretical understanding of stellar flares, and builds context for the rest of the thesis. The primary dataset TESS will be introduced in more detail during chapter 3. This chapter also covers the stars chosen for analysis in this thesis, and the previous work on mapping their surface activity using Doppler imaging and other methods.

Chapter 4 describes the numerical methods and approaches used in this work. Several methods from recent literature are used on the dataset. The final chapter explores the results of the different methods, and discusses their implications and applicability for future research.

2. Theory

2.1 Active stars

2.2 Magnetohydrodynamics

2.3 Flares

3. Datasets

3.1 TESS mission

3.2 The stars

3.3 Preprocessing and validation

4. Numerical methods

4.1 Periodic analysis

4.2 Detection

4.2.1 FLATW'RM

A 2018 article outlines a machine learning method to detect flares from periodic signals, named FLATW'RM ("FLAre deTection With Ransac Method"). [Vida and Roettenbacher, 2018] This approach is going to be particularly useful for the purposes of this thesis, as its approach is robust enough to fit even a multiple periodic signal.

Its backbone is the RANSAC ("Random Sample Consensus") algorithm introduced in the Communications of ACM journal. [Fischler and Bolles, 1981] It is applied to find a robust fit of the lightcurve in a window of the lightcurve. An outline of this method follows. Let n be the minimum number of datapoints to fit a given model to the signal, S the signal ($|S| > n$), and m a threshold of points that we require to lie in a desired confidence interval from the model such that $|S| \geq m > n$.

1. At random, pick a sample $S' \in S$ of size n.
2. Compute the fit M' using S'.
3. Let S'' be the portion of S that falls within the desired confidence interval of M'.
4. If $|S''| \geq m$, compute and accept the new fit M'' based on S''.
5. If not, go to 1.

This randomized sample consensus effectively prevents a few outliers from dominating the fit, by rejecting models that produce a large portion of outliers. Its complexity is proportional to the complexity of the fit and the inverse probability of the fit being accepted. We can improve RANSAC by adding further tolerances $m < m' < m'' < ...$ that let us iterate steps 2-4 to create an even more robust fit.

FLATW'RM goes through the data series in a rolling window, and fits a model for each window with RANSAC. The size of the window is around one primary period of the lightcurve. Each window picks flare candidates based on outliers, similar to ordinary models. After the window has rolled through the signal, the windows vote on the proper flare candidate.

5. Results and Discussion

Bibliography

T. Barclay, J. Pepper, and E. V. Quintana. A Revised Exoplanet Yield from the Transiting Exoplanet Survey Satellite (TESS). , 239(1):2, Nov. 2018. doi: 10.3847/1538-4365/aae3e9.

J. R. A. Davenport. The Kepler catalog of stellar flares. *The Astrophysical Journal*, 829(1):23, Sep 2016. ISSN 1538-4357. doi: 10.3847/0004-637x/829/1/23. URL `http://dx.doi.org/10.3847/0004-637X/829/1/23`.

T. V. Doorsselaere, H. Shariati, and J. Debosscher. Stellar Flares Observed in Long-cadence Data from the Kepler Mission. *The Astrophysical Journal Supplement Series*, 232(2):26, Oct 2017. ISSN 1538-4365. doi: 10.3847/1538-4365/aa8f9a. URL `http://dx.doi.org/10.3847/1538-4365/aa8f9a`.

M. A. Fischler and R. C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Commun. ACM*, 24(6):381–395, June 1981. ISSN 0001-0782. doi: 10.1145/358669.358692. URL `https://doi.org/10.1145/358669.358692`.

S. T. e. a. Maehara, Hiroyuki. Superflares on solar-type stars. *Nature*, 485:478–481, 2012.

G. R. Ricker. The Transiting Exoplanet Survey Satellite Mission. *Journal of the American Association of Variable Star Observers (JAAVSO)*, 42(1):234, June 2014.

E. Rubenstein, Ph. D. and B. Schaefer. Are Superflares on Solar Analogues Caused by Extrasolar Planets? *The Astrophysical Journal*, 529:1031, 12 2008. doi: 10.1086/308326.

B. E. Schaefer, J. R. King, and C. P. Deliyannis. Superflares on Ordinary Solar-Type Stars. *The Astrophysical Journal*, 529(2):1026–1030, Feb 2000. ISSN 1538-4357. doi: 10.1086/308325. URL `http://dx.doi.org/10.1086/308325`.

Z.-L. Tu, M. Yang, Z. J. Zhang, and F. Y. Wang. Superflares on Solar-type Stars from the First Year Observation of TESS. *The Astrophysical Journal*, 890(1):46, Feb 2020. ISSN 1538-4357. doi: 10.3847/1538-4357/ab6606. URL `http://dx.doi.org/10.3847/1538-4357/ab6606`.

K. Vida and R. M. Roettenbacher. Finding flares in Kepler data using machine-learning tools. *Astronomy Astrophysics*, 616:A163, Aug 2018. ISSN 1432-0746. doi: 10.1051/0004-6361/201833194. URL `http://dx.doi.org/10.1051/0004-6361/201833194`.