Stellar Flares Observed in Long-cadence Data from the *Kepler* Mission

Tom Van Doorsselaere[1] , Hoda Shariati[2], and Jonas Debosscher[2,3]

[1] Centre for mathematical Plasma Astrophysics, Mathematics Department, KU Leuven, Celestijnenlaan 200B bus 2400,
B-3001 Leuven, Belgium; tom.vandoorsselaere@kuleuven.be

[2] Instituut voor Sterrenkunde, KU Leuven, Celestijnenlaan 200D, B-3001 Leuven, Belgium

[3] Royal Belgian Institute for Space Aeronomy, Ringlaan-3-Avenue Circulaire, B-1180 Brussels, Belgium

Abstract

We aim to perform a statistical study of stellar flares observed by *Kepler*. We want to study the flare amplitude, duration, energy, and occurrence rates, and how they are related to the spectral type and rotation period. To that end, we have developed an automated flare detection and characterization algorithm. We have harvested the stellar parameters from the *Kepler* input catalog and the rotation periods from McQuillan et al. We find several new candidate A stars showing flaring activity. Moreover, we find 653 giants with flares. From the statistical distribution of flare properties, we find that the flare amplitude distribution has a similar behavior between F+G types and K+M types. The flare duration and flare energy seem to be grouped between G+K+M types versus F types and giants. We also detect a tail of stars with high flare occurrence rates across all spectral types (but most prominent in the late spectral types), and this is compatible with the existence of "flare stars." Finally, we have found a strong correlation of the flare occurrence rate and the flare amplitude with the stellar rotation period: a quickly rotating star is more likely to flare often and has a higher chance of generating large flares.

Key words: stars: activity – stars: coronae – stars: flare

Supporting material: figure set, machine-readable table

1. Introduction

It is well known that not only the Sun is showing magnetic activity in the form of flares (e.g., Kowalski et al. 2010; Hawley et al. 2014). On the Sun, it is believed that solar flares are caused by magnetic reconnection (e.g., Sun et al. 2015). Thus, it is a straightforward assumption that stellar flares are also caused by magnetic reconnection of coronal structures. Hence, the presence of flares can be used as a proxy for the presence of a stellar corona.

Before space photometry, stellar flare studies were focused on active flare stars, in order to increase the chance of capturing a flare during the limited telescope observing time. Some early statistical studies exist (Shakhovskaia 1989; Kowalski et al. 2013), but they are rare. However, with the *Kepler* mission, it became possible to systematically study stellar flares. The first to use the *Kepler* data for this purpose were Walkowicz et al. (2011), and there they limited the sample to cool dwarfs. To study superflares, Maehara et al. (2012, 2015), Shibayama et al. (2013), and Candelaresi et al. (2014) selected a sample of G-type stars or later, to obtain power laws for flare amplitudes per star. Moreover, several case studies of flare stars using *Kepler* data have been published (e.g., Ramsay et al. 2013; Lurie et al. 2015), and those studies are continued with K2 as well (Ramsay & Doyle 2014).

Balona (2012, 2013) used visual inspection of the light curves to identify flares in a much broader (in terms of spectral type) sample of *Kepler* observations. They found that even some A stars show flaring activity. This is unexpected from stellar evolution theory, because these stars are not believed to have an outer convection zone. Due to the dynamo effect, the latter is considered to be a crucial ingredient for the star to show magnetic activity, such as flares. Thus, it is plausible to ascribe the flares on A stars to cool companions. However, it was argued by Balona (2012, 2015) that this could not be the case, because then the flare amplitude would be unusually large for these cool companions. Still, Pedersen et al. (2017) studied the list of A stars of Balona (2013) in great detail. They found that several cases could be explained by contaminated pixel data or found that several of the flaring A stars were in a binary (implying that the flare is originating from the companion). Yet, not all flaring A stars could be excluded, and thus it remains inconclusive if A stars can flare or not.

More recently, Pitkin et al. (2014) and Davenport (2016) developed automated algorithms to process the large *Kepler* database and study flares in a statistical sense. Still, several aspects are missing from those studies. For example, it was previously shown that X-ray luminosity (presumably from the stellar corona) scales with the rotation period of the star (Wright et al. 2011), but this aspect was not studied in the previous statistical studies on stellar flares.

With the high-quality data from *Kepler*, even seismology of stellar flares can be attempted using quasi-periodic pulsations in stellar flares (for a review, see Van Doorsselaere et al. 2016). Observations of stellar flare oscillations were reported by Mathioudakis et al. (2003), Mitra-Kraev et al. (2005), Welsh et al. (2006), Anfinogentov et al. (2013), Srivastava et al. (2013), and Balona et al. (2015), with even multiperiodic events (Pugh et al. 2015) being detected now. Given the similarity with quasi-periodic pulsations in solar flares (Cho et al. 2016), a lot of potential exists for remote sensing of stellar coronae and their magnetic fields.

In this paper, we develop a new automated detection and characterization method for flares in the *Kepler* mission data. We perform a statistical study of the stellar flares. In particular, we study the dependence of the flare occurrence rate, flare duration, and flare amplitude on stellar spectral type and stellar rotation periods.

2. Detection Algorithm

2.1. Preprocessing

We use the raw light curves from the *Kepler* mission during quarter 15 (Q15 for short). We focus on the long cadence, which has a time cadence of approximately $\Delta t = 29$ minutes. Each time series consists of different segments between data gaps and intensity discontinuities, which are instrumental in nature. For each of these segments, we fit a third-order polynomial to remove the instrumental effects. We have removed the detected flare candidates near these discontinuities, to ensure that our flare sample is not influenced by these instrumental effects.

After this detrending, we have prewhitened the time series, following the procedure outlined in Degroote et al. (2009). We have computed the Lomb–Scargle periodogram of the data to determine the most significant frequency peak in the interval $f = [0.1 \ \mathrm{day}^{-1}, 24.5 \ \mathrm{day}^{-1}]$. If this peak is statistically significant (with S/N > 4 using the white noise approximation), then a sine with this frequency is fitted to the data, and then removed. This procedure is repeated until no significant peaks are found, or at most 100 frequencies have been removed from the light curve. This prewhitening procedure removes most of the regular periodic effects, such as intensity variations caused by stellar pulsations or stellar spots. However, the prewhitening does not do so well for eclipsing binaries, and thus any flare events on known eclipsing binaries are removed from the database (using the information at http://keplerebs.villanova.edu).

2.2. Thresholding

The detrended and prewhitened light curve is then subjected to a double threshold method, as illustrated for an example flare in Figure 1. We have first computed the standard deviation of the central difference of the time series Δf $((f_{i+1} - f_{i-1})/2$ at time t_i, which estimates its slope). After making an initial estimate of the standard deviation $\sigma_{\Delta f}$, we have made this estimate more robust by removing once all outliers above $6\sigma_{\Delta f}$ from the central difference time series.

Then we check if the flare intensity is sufficiently higher than the mean intensity of the star: the threshold for detection of a flare is set at $4.5\sigma_{\Delta f}$, which statistically corresponds to $4.5\sigma_f/\sqrt{2}$, where σ_f is the standard deviation of the noise of the detrended and prewhitened flux $f(t)$.

The second threshold is in the central difference of the flux $\Delta f(t)$, in order to check that the increase in intensity was sufficiently rapid, following the FRED profile (fast rise exponential decay). The running difference threshold is taken as $3\sigma_{\Delta f}$. Additionally, we check if an interval of four data points on either side of the flare maximum at time t_0 contains a maximum in the flare slope left of the peak ($[t_0 - 4\Delta t, t_0]$), and a minimum in the slope right of the peak ($[t_0, t_0 + 4\Delta t]$). This last criterion ensures that the flux changes are rapid enough.

2.3. Parameterization

After the detection, we perform further filtration of the detected flare candidates. We want to study decay times of flares, and therefore we fit the detected flares with an exponentially decaying function, following Anfinogentov
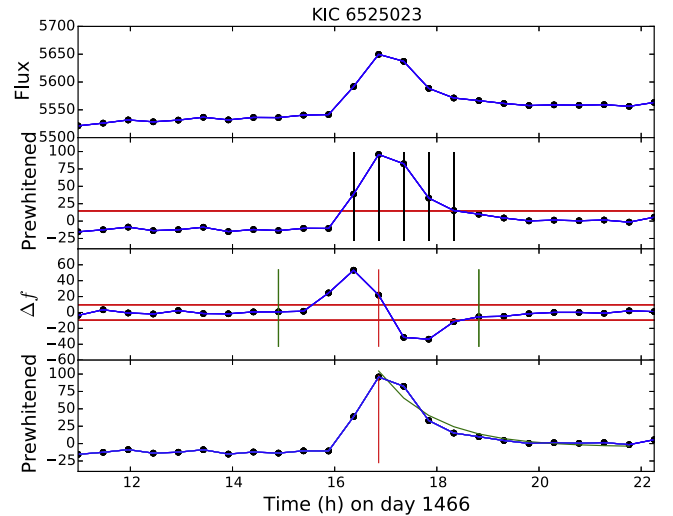

Figure 1. Illustration of the detection method. The top panel shows the raw data from the *Kepler* database (black dots) for KIC 6525023 as a function of time on day 1466 of the *Kepler* mission. The second panel shows the detrended light curve $f(t)$ (black dots) with the employed threshold (red line). The points above the threshold are highlighted with vertical black lines. The third panel shows the central difference Δf of the intensity signal (black dots), the thresholds (horizontal red lines), and the flare window (vertical green lines). The bottom panel shows the detrended light curve (black dots), with the flare peak indicated (vertical red line), and the fitted exponential decay (green line).

et al. (2013) and Pugh et al. (2016):

$$g(t - t_0) = a \exp(-k(t - t_0)) + b. \qquad (1)$$

To do this, we introduce three more constraints on the flare detection. First, we determine the length of the time series to be fitted. The start of the time series is the time of the peak of the flare t_0. The time t_e is found as the time for which the gradient Δf is closest to zero, and it has to be within the five points right of the last point where the flux f exceeds the threshold. For the fitting, four more data points are added to the time series, and the fit is thus performed in the time interval $[t_0, t_e + 4\Delta t]$.

Then, there are three possible reasons to reject the flare: (1) there are multiple local maxima in flux during the flare in the time interval $[t_0, t_e]$ and a FRED profile is thus not a good match, or (2) the fit with the exponential function g fails, or (3) the flare has a negative amplitude $a \leqslant 0$ or the amplitude is smaller than the background $a \leqslant b$.

It is worth noting that there is no restriction on the duration of the flare $1/k$, and thus also short-duration flares ($1/k < 30$ minute) are retained in the results. However, we have checked the results in Section 3 by excluding those short flares. We find that all incidence rates are decreased by 0.50%, except for the giants (where the incidence rate remains nearly constant). This tells us that the short-duration flares are uniformly distributed over all spectral types, except for the giants. Moreover, the distributions of flare amplitudes and energies (see Section 3.5) are not modified by the exclusion of short-duration flares.

3. Results

3.1. General Interpretation

We have constructed a Hertzsprung–Russell (HR) diagram in Figure 2. To this end, we have used the information in the *Kepler* Input Catalog (KIC, v10 downloaded from
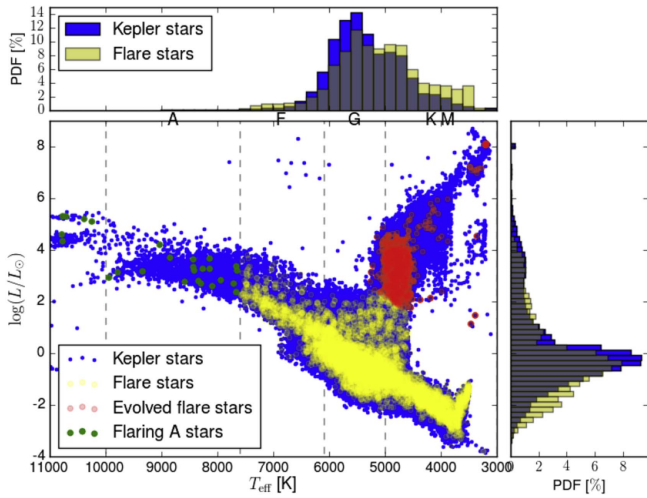
Figure 2. HR diagram with the detected flare stars indicated. The central panel has stellar temperature as the horizontal axis, while the vertical axis shows the luminosity. *Kepler* stars are shown with blue dots. Yellow, green, and red dots indicate stars with detected flares, yellow dots are close to the main sequence, green dots are A stars with flares, and red dots are giants. The top and side panels show the histograms of the *Kepler* stars (in blue) and flare stars (in yellow).

https://archive.stsci.edu/kepler/kic10/search.php), where we have extracted the temperature $T_{\rm eff}$ and used the stellar radius $R/R_{\odot}$ to estimate the luminosity L (in units of the solar luminosity $L_{\odot}$) by

$$L = \ln\left\{ \left(\frac{R}{R_{\odot}}\right)^2 \left(\frac{T_{\rm eff}}{5780\ \mathrm{K}}\right)^4 \right\}. \qquad (2)$$

To estimate the spectral type, we have taken an interval of [3000 K, 5000 K] as K and M stellar types, [5000 K, 6100 K] as G stars, [6100 K, 7600 K] as F stars, and [7600 K, 10,000 K] as A stars. These temperature intervals are shown by vertical dashed lines in Figure 2.

The stars and detected flares are listed in Table 1. The table contains a line for each detected flare candidate with the Kepler ID of the object, flare start time, flare amplitude a, flare decay k, the effective temperature $T_{\rm eff}$, stellar radius $R/R_{\odot}$, and $\log g$.

In Figure 2, the *Kepler* stars are shown with blue dots, while the flare stars are shown with red, green, or yellow dots. The top and right panels show the normalized histograms of the distribution of *Kepler* and flare stars over temperature and luminosity. It is clear that the flare distribution (in yellow) is more concentrated toward the lower end of luminosity and temperature. This confirms the finding of Davenport (2016) that the flare star occurrence rate increases toward later spectral types.

In total, we have detected 16,850 flares on 6662 stars out of a total of 188,837 in the *Kepler* field of view during Q15. This translates into 3.5% of the stars being flare stars. This number is higher than the percentages mentioned in Walkowicz et al. (2011), Balona (2012), and Davenport (2016). Especially the latter work is very relevant, because he also uses an automated detection method as we do. However, Davenport (2016) has removed all flare star candidates with less than 25 flares during the entire *Kepler* observations. We have kept all flare stars, even if they only show one flare, because we want to avoid a selection effect in the sample.

In Table 2, we give an overview of the flare star incidence for each spectral type. A star can be counted both as a giant and as belonging to its spectral type. These incidence rates are relatively close to the values found in Davenport (2016). Also in this table, it is confirmed that the flare star incidence is twice as high for K and M stellar types, compared to F and G stellar types. Our flare star incidences seem to be different from the statistics from Balona (2015): he found a much higher incidence rate of 12% for K- and M-type stars, and a reduced incidence rate of 1.2% for F-type stars, but did find a similar incidence rate for the G-type stars. The differences between our study and that of Balona (2015) could be caused by the much smaller sample size of the latter work.

3.2. False Detections and Detection Errors

It is important to quantify the number of false detections and detection errors of our algorithm, because this will likely influence the statistics of the stellar flare detections. The reasons for false detections could be twofold: (1) the algorithm may not function as expected or (2) there are signals coming from nearby stars that fall in the *Kepler* pixel mask. Note that the first was circumvented by Davenport (2016) by excluding stars with fewer than 25 flares, while the latter was avoided by Maehara et al. (2012) by eliminating stars that have another star within 12 arcsec of the target.

To assess the false detection or detection errors, we have performed Monte Carlo simulations. To do this, we have first fixed a value for the possible false-detection probability $p_{\rm false}$ (with values between 0.1% and 30%; see horizontal axis of Figure 3). Then, we have randomly chosen $p_{\rm false} N_{\rm flare\ star}$ out of the fixed number of detected flare stars $N_{\rm flare\ star}$. For each of these randomly chosen flare stars, we have replaced their characterization from the KIC by a random value from another star in the field. This is then mimicking the fact that the flare could occur on a background star (within the pixel mask), for which the spectral type, temperature, and radius distribution are the same as the distribution of *Kepler* (nonflaring) stars. Then, we have repeated this process 1000–3000 (nonflaring and flaring) times for a different random selection of stars (using the Monte Carlo spirit).

Figure 3 then shows the mean (with blue dots) of the obtained distribution for the fraction of flaring A stars, and the 1σ error bars (given by the square root of the variance of the distribution). The green dotted lines show the 2σ interval around each mean. It is clear that the mean is increasing for a larger false-detection probability $p_{\rm false}$, and this is expected, because the mean should evolve to the global mean (over all spectral types). In fact, when $p_{\rm false} = 1$, it should almost go to the global average of 3.5% of flare stars in the population of stars. The red dashed line (in Figure 3) shows the fraction of flaring A stars that were found for the total population (see Table 2). Since it makes physical sense that the fraction is smaller than for the average field stars (composed mostly of G types), it is probably a good assumption that the obtained value of 1.31% must fall within the 2σ uncertainty strip of the Monte Carlo simulation. This allows us to estimate our false-detection probability as being at most 15%.

3.3. Active A Stars

As described in the introduction, flares were surprisingly found on A stars by Balona (2012). In Figure 2, it is clear that

Table 1
Catalog of *Kepler* Flare Stars

KIC	Date	a	k	T_{eff}	$R/R_{\odot}$	log g
757099	1379.761379	1537.188432	29.809984	5589.0	2.288	3.817
757099	1384.174846	501.038611	24.993707	5589.0	2.288	3.817
757099	1388.996955	439.617457	49.042552	5589.0	2.288	3.817
757099	1393.941649	396.729576	76.259085	5589.0	2.288	3.817
757099	1410.655473	1037.415398	27.823375	5589.0	2.288	3.817
757099	1434.643222	405.202566	34.496090	5589.0	2.288	3.817
757099	1454.646748	439.380565	17.406161	5589.0	2.288	3.817
757099	1458.222475	602.339598	27.006658	5589.0	2.288	3.817
757450	1449.436419	24.900689	29.209920	5101.0	0.933	4.48
892376	1387.362344	149.690763	8.227514	3813.0	0.7	4.471
892376	1401.297377	362.261581	13.862022	3813.0	0.7	4.471
892376	1446.289793	155.572817	26.897590	3813.0	0.7	4.471
892376	1465.578286	88.628808	31.918374	3813.0	0.7	4.471
893507	1373.774593	1149.538801	13.145064	5212.0	1.835	3.978
893507	1386.361162	526.195450	11.850842	5212.0	1.835	3.978

(This table is available in its entirety in machine-readable form.)

Table 2
Flare Star Incidence over the HR Diagram

Stellar Type	# Objects	# Flare Stars	Incidence
A+B	2141	28	1.31%
F	22107	708	3.20%
G	116178	3365	2.90%
K+M	48411	2556	5.28%
giants	22837	653	2.86%

Notes. The first column indicates the stellar type, the second column indicates the number of objects in this type, the third column shows how many have at least one flare, and the fourth column is the ratio of Column 3 by Column 2 to get the incidence of flare stars for the given type.

Figure 3. Fraction of A stars that show flaring as a function of the false-detection probability. The blue dots show the mean fraction of the Monte Carlo results, while the blue bars show the 1σ errors. The green dotted lines show the 2σ uncertainty strip. The red dashed line shows the value with a false-detection probability of zero (i.e., the algorithm is perfect) and is also given in Table 2.

some flares were found in stars with temperature hotter than 7600 K, which we classify as A stars (indicated with green dots). The stellar objects and their effective temperatures are listed in Table 3. As can be seen, some temperatures even go higher than 10,000 K (and are possibly B or even O stars), but given the uncertainty and bias on temperatures in the data from the KIC, it is safer to assume that these are A stars as well. The table lists in the right column possible alternative explanations for the flares, as given by Pedersen et al. (2017). However, our study finds 24 new A stars that show flaring activity. Thus, we add to the body of evidence that there may be some A-type stars that show magnetic activity.

In Appendix B, we show the flare light curves for the A stars. An uncritical look will reveal Figures 9.2, 9.13, 9.22, 9.29, and 9.30 of the Figure 9 figure set as being unreliable flares (indicated as "Unreliable" in the third column in Table 3). Furthermore, an even more critical look will also cast doubt on the flares in Figures 9.3, 9.10, 9.11, and 9.19, because the stellar variability has the same timescales as the detected flares (indicated with "Doubt" in the third column in Table 3).

Taking only the "unreliable" flare stars as misdetections, we have misdetected 10 flares out of 61 flares, amounting to 16% of the detected flares. With this assumption, the false-detection probability of flaring A stars is thus 4 out of 28, which is 14%. This value is close to our misdetection probability as estimated
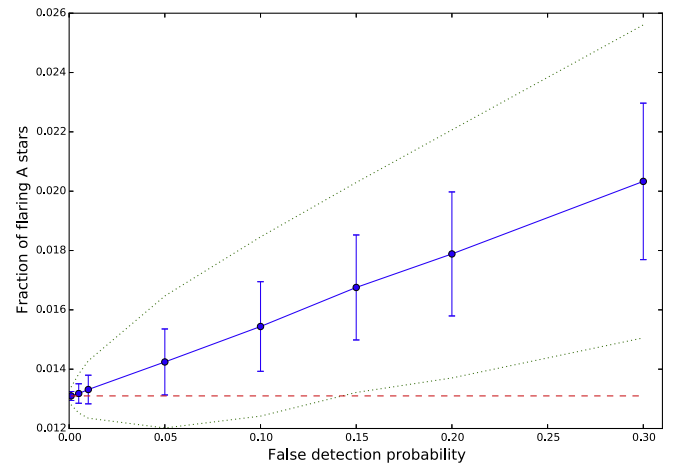
in Section 3.2. However, when also taking into account the A stars in "doubt," the misdetection probability goes as high as 25%.

One may also argue that the A stars disqualified by Pedersen et al. (2017) should not be listed in Table 3. However, the automated detection algorithm *should* detect these stars: indeed, they appear as A stars with flares and are thus correctly detected by our algorithm (despite the flares most likely originating on a cool companion or neighbor; Pedersen et al. 2017).

3.4. Flaring Giants

Most of the flaring stars are found on (or near) the main sequence. On the other hand, several objects have a temperature and luminosity located in the red giant branch or even the asymptotic giant branch. A few such objects were also previously detected by Balona (2015). These giant stars with

Table 3
List of A Stars with a Flare Detected

Kepler ID	$T_{\rm eff}$ (K)	Comment	Pedersen et al. (2017)
1294756	8411.0	New	
1430353	10765.0	New	
		Unreliable	
4547333	10769.0	New	
		Doubt	
4573879	8288.0	New	
5113797	8139.0		Overlapping neighbor?
5201872	7937.0		Cool companion
			Overlapping neighbor?
5273195	8588.0	New	
5284647	9777.0	New	
5632093	8085.0	New	
5879187	7668.0	New	
		Doubt	
5905878	8337.0	New	
6954726	16764.0	New	
		Unreliable	
7097723	8438.0		Contamination
			Cool companion?
7523115	8438.0	New	
7599132	10251.0	New	
8044889	7653.0		Cool companion
8129631	9946.0	New	
		Doubt	
8142623	9332.0	New	
8214398	8848.0	New	
8264075	7654.0	New	
		Unreliable	
8515910	8143.0	New	
8881883	10710.0	New	
10149211	10785.0	New	
10593239	8259.0	New	
10974032	9038.0	New	
11293898	15072.0	New	
		Unreliable	
11912716	10386.0	New	
11919968	7707.0	New	

Notes. The left column shows the stellar Kepler ID, the middle column shows the stellar temperature $T_{\rm eff}$, and the third column indicates which objects are new here, but also lists unreliable flares (see Section 3.3). The right column indicates the explanations as listed in Pedersen et al. (2017) for objects that were previously found.

flares are shown by red dots in Figure 2. To be classified as giants, the flare stars have to satisfy the following relation:

$$\frac{L-1}{1.5} > \frac{T_{\rm eff} - 4200\ {\rm K}}{1000\ {\rm K}}. \qquad (3)$$

This relation was inspired by the shape of the HR diagram and serves to separate the giant branch from the main sequence. In total, we find 653 giants with flaring activity.

As for A stars, it is equally unexpected that giant stars have strong magnetic fields. During their evolution on the main sequence, stars spin down. Moreover, their increase in size would lead to a decreased surface magnetic field (because of the magnetic field dependence on distance). As a result, their magnetic field is expected to be weaker (Simon & Drake 1989).

This was confirmed observationally by Konstantinova-Antova et al. (2008), who found a magnetic field of 5 G and 15 G for a rapidly rotating giant. However, outbursts that were attributed to stellar activity were observed on Mira A by Karovska et al. (2005) and Vlemmings et al. (2015). Harper et al. (2013) found chromospheric emission on giants, and Gaulme et al. (2014) found light curve variations compatible with stellar spots on giants.

Given the numerous evolved objects with flares revealed by our analysis, it becomes clear that magnetic activity is ubiquitous on the red giant branch. Moreover, the flare star incidence for giants (as given in Table 2) is similar to that of F and G stars. This value of the flare star incidence among giants may be compatible with two scenarios for the explanation of flares on these stars: (1) The flares are actually attributable to a companion of K or M type. The incidence rate among giants is around half of that of K and M stars, which would then require a cool companion to half of the flaring giants. (2) Alternatively, one could argue that this value of the flare star incidence among giants is compatible with the flare star incidence in their progenitors, i.e., the F and G stars. This could mean that the magnetic activity of the stellar atmosphere is more or less conserved during its evolution from the main sequence to the red giant branch.

One may also wonder about the nature of the flares. Perhaps these flares are not caused by the same mechanism as on late-type main-sequence stars, where it is believed an internal dynamo leads to complex surface magnetic fields, reconnection, and flares. The giants may have a different mechanism that operates to create these outbursts. There may be an enhanced surface dynamo (e.g., Amari et al. 2015), or the extended magnetic loops could reconnect due to centrifugal forces.

3.5. Flare Amplitudes

In Figure 4, we show the (logarithm base 10 of the) histograms of the flare amplitude and energy for different stellar categories. The absolute amplitudes (middle panels) are calculated by the flare amplitude minus the preflare stellar intensity, and the relative amplitude normalizes the absolute flare amplitude by the preflare stellar intensity. The flare energy $E_{\rm flare}$ is calculated by using Equation S5 from Maehara et al. (2012, equivalent to Equation (6) of Shibayama et al. 2013, when assuming the exponential intensity profile of Equation (1)):

$$E_{\rm flare} = \frac{a}{k}\exp(L)L_{\odot}, \qquad (4)$$

with $L_{\odot}$ given in erg, and its distribution is shown in the right panels.

The total population of flaring stars is subdivided into spectral types using the temperatures listed in the KIC (as explained in Section 3.1). We do not show the results of A-type stars or hotter, because there is only a limited number of stars in that category (see Section 3.3). Moreover, the K+M types (third row) have been split between giant stars and dwarf stars (using Equation (3)).

The horizontal scales of the histogram have been limited to 10% and 3000 e$^-$ s^{-1}, because some flares with much higher amplitudes have been detected, but these events disallow a
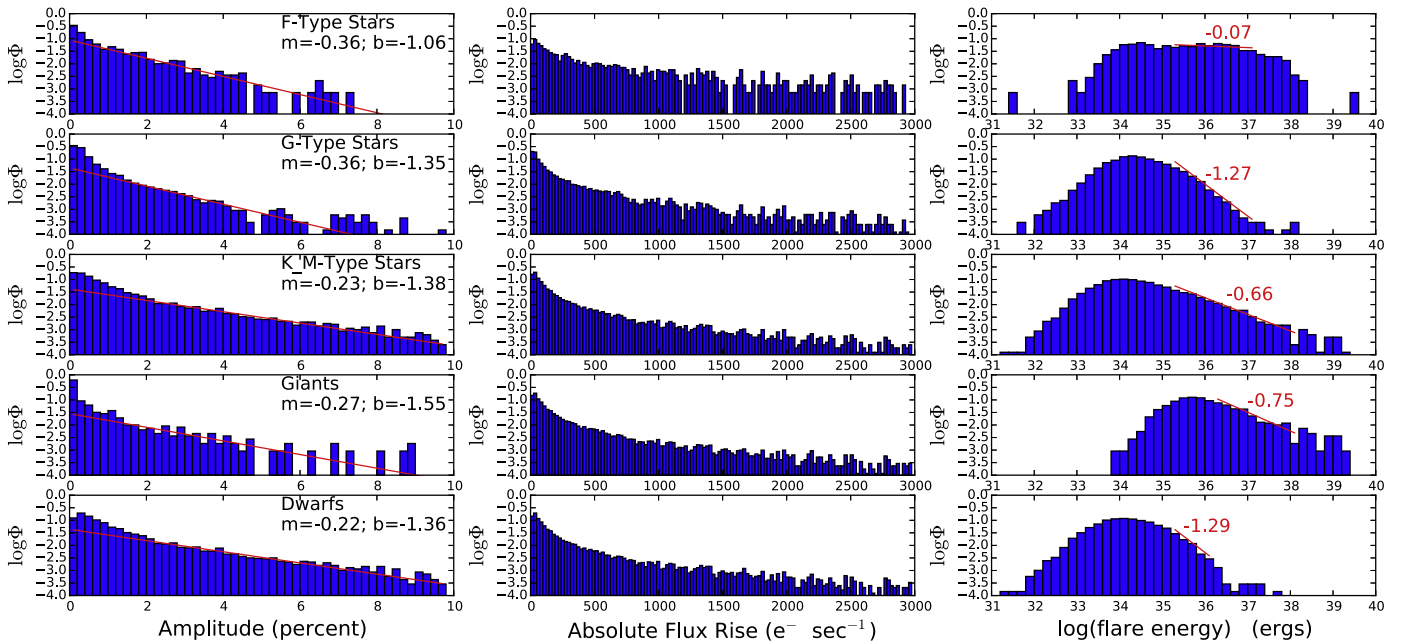
Figure 4. Histogram of flare amplitude and energy per spectral type. The vertical axis is the logarithm of the histogram count. The horizontal axis is the relative flux increase in the left panel, the absolute flux increase in the middle panel, and the logarithm of the flare energy (in erg) in the right panel. The left panel shows the exponential fit between 2% and 4% with a red line, while the right panel shows the fit only in the fitting range.

clear comparison between the different spectral types. The histogram of the relative amplitude has been fitted with an exponential distribution (10^{mx}), motivated by its appearance. This is done by fitting a straight line $y = mx + b$ (with y the logarithm base 10 of the histogram and x the relative amplitudes) by a regression between the amplitudes $x = 2\%$ and $x = 4\%$ to avoid the influence of the tail (which is presumably dominated by small-number statistics). The fitted values of m and b are shown in the key.

From the histograms and their fits, it is clear that the F- and G-type stars show the same behavior in the relative amplitude, and even an identical slope is found from the fit. This is an indication that the magnetic phenomena on these types of stars have the same underlying physical mechanism, because the relative flare amplitude is equally distributed. However, for the flare energy, a widely different behavior is found. The F stars show a flat energy distribution, while the G-type stars show a steeply decreasing energy distribution.

For the K- and M-type stars, the histogram for the relative amplitude seems to be different. There seem to be more flares with a high amplitude, and they form a longer tail in the histogram. This is also confirmed by the slope fit of -0.23, which is shallower than the early-type stars (F and G). We have also split the K- and M-type stars between giant and dwarf stars and show their histograms separately. Given the smaller number of flaring giants, the distribution of the amplitudes of flares on dwarfs is very close to the distribution of the K- and M-type stars. Even the fitted slope is not changed very much. However, the distribution of the amplitudes of flares on giants is quite different from the distribution of the dwarf stars, because the heavy tail of large-amplitude flares is missing from the distribution. This is also confirmed by the fitted slope: it is steeper than the slope of the dwarf amplitude distribution.

From these results, we may tentatively conclude that the magnetic mechanism for flares is the same on F- and G-type stars, but differs for dwarf stars (K and M types). For giants, it is unclear what magnetic mechanism they could use to generate flares. On the one hand, it could be that there is yet another mechanism at work to generate flares on giants. On the other hand, it could be that there are two mixed populations of giants, with each population generating flares with the magnetic mechanism of F and G types or that of the dwarfs. One of the two populations could be constituted of giants that were already flaring during their life on the main sequence, and the other population may generate flares with a mechanism similar to that of the dwarf stars.

From the flare energy distributions (right panel, Figure 4), we see that the low-energy end is strongly influenced by observational bias, because only the stronger flares are detected. It also shows that the flare detection routine (which is based on thresholds) is influenced by the background brightness of the star, because the peak of the flare energy distribution depends on the stellar type. In particular, the flare energy distribution for the giants (and to a lesser extent for the F stars) is shifted toward higher energies. This is because only the very strongest flares are detected against the much stronger emission from the giants (due to their larger radius).

The high-energy tail is governed by an apparent power law, which is physical in nature. The fits to the power law (with 10^{mx}) are shown with red lines in the fitting regions, and their slopes are indicated. The F-type stars have a very shallow distribution, but the power-law slope of the G types and dwarfs is consistent. The giants have a slope between that of the flat F-type distribution and the G-type and K+M-type distributions. As we have inferred from the amplitudes as well, it could point in the direction of two flare-generation mechanisms at work in flaring giant stars.
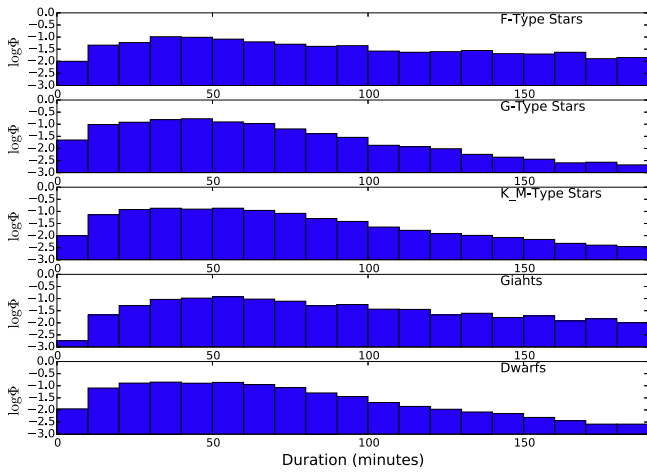
Figure 5. Histogram of flare duration per spectral type. The vertical axis is the logarithm of the histogram count. The horizontal axis is the duration of the flare in minutes, taken as $1/k$ (Equation (1)).

Figure 6. Histogram of flare occurrence rate per spectral type. The vertical axis is the logarithm of the histogram count. The horizontal axis is the flare occurrence rate in units of flares per star per quarter. The key contains the fitting parameters of a straight line (shown in red) and the MLE values for the exponential parameter (shown in olive).

3.6. Flare Duration

In Figure 5, we show the histograms for the duration of the flares, split between stellar types as in Section 3.5. The duration is calculated as $1/k$ from the fit with Equation (1). It is clear that the histograms do not follow an exponential distribution, as seemed to be the case for the amplitude distribution (Figure 4). The reason is clear: we have used long-cadence data from the *Kepler* mission, and thus we have only a small chance of observing flares with a duration shorter than 30 minutes. Moreover, flare detections of only one data point have been rejected, further limiting our detection capabilities in that range. Thus, the location of the peak of the distribution is mainly determined by the cadence time of the telescope and our detection algorithm. Still, it seems that the peak of the histogram for the duration of flares on giants is shifted to the right with respect to the other histograms. One may wonder if this fact contains any physics.

From the graph, it seems that now the G- and dwarf K+M-type stars have a similar behavior for the flare duration. This is surprising, because the previous section found that the amplitudes of flares on G stars behave more like the amplitudes of flares on F stars. However, now it is clear that the flares last much longer on F stars than on G stars. So, it seems that the separation between the two flare generation mechanisms that we postulated in Section 3.5 is not so clear-cut. It could be that the transition between the two mechanisms does not coincide with the chosen boundary of spectral types, or that there is an overlapping region in the HR diagram where the two flare mechanisms occur simultaneously.

The distribution of duration of flares on giants follows closely the distribution of F-type stars, with a heavy tail toward longer duration. Thus flares on giants have statistically a longer duration than flares on dwarf K+M stars. This could point in the direction that their flaring mechanism has remained the same as during their main-sequence life as an F star. However, such a conclusion is not supported by the difference between F stars and giants in Figure 4.

The measured duration of the flares may be heavily influenced by the thresholds we chose for the detection. So, some words of caution should be added that the results of this
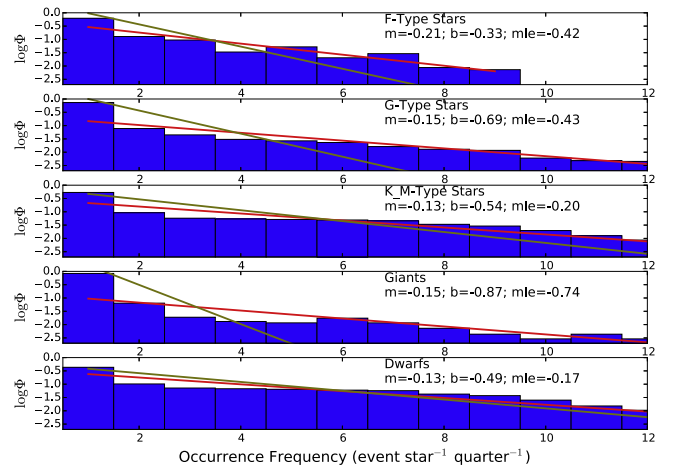
section may change for different detection algorithms, and this could be tested by other teams (e.g., Davenport 2016).

3.7. Flare Occurrence Rate

Figure 6 shows the histogram for the occurrence rates of flares per object. Since most of the histograms seem close to an exponential distribution (10^{mx}, x is the occurrence rate here), we have fitted the histogram with a straight line $y = mx + b$ (with y the logarithm of the histogram counts). The parameters of this fit have been displayed in the key on each graph. We have also fitted the distributions by using the maximum likelihood estimator (MLE) for m, assuming that the offset-exponential distribution starts from a value θ. Following Johnson & Kotz (1970), the MLE estimator for the shifted exponential distribution is given by

$$\mathrm{MLE}_\theta = \frac{n}{\sum_{i=1}^{n} x_i - n\theta} = \frac{1}{\mu_S - \theta}, \qquad (5)$$

with $S = \{x_i | i = 1, \ldots, n\}$ the n samples of the data, and μ_S is the sample's mean. This is analogous to the width estimator of the Laplace distribution.

Thus, $\mathrm{MLE}_1 \cdot (\log_{10} e)$ is the MLE for m, assuming a minimum occurrence rate of one event per star per quarter, because we did not include stars without flares. In the figure, we have printed the obtained value $\mathrm{MLE}_1 \cdot (\log_{10} e)$, which is the estimator for m. Any further mentions of MLE values always include the $\log_{10} e$ factor, to relate to the 10^{mx} proposed distribution.

On the one hand, the MLE_1 value is dominated by the bulk of the distribution and therefore contains information on the stars that do not flare very often. On the other hand, the fitting is more heavily influenced by the tail of the distribution and thus contains information about the most frequently flaring stars.

From the MLE values in the keys of Figure 6, it seems that the K+M-type stars have a higher flare occurrence rate, followed by F and G, and that the giants have the lowest flare occurrence rate. If K+M stars are flaring, they flare on average 3.1 times per quarter, flaring F and G stars are flaring on

Table 4
Overview of Flare Star Incidence in the Database of McQuillan et al. (2014b)

log Ro	N_{stars}	N_{flare}	%
−2.5 to −1.9	26	18	69.2 ± 10.9%
−1.9 to −1.6	174	88	50.6 ± 5.3%
−1.6 to −1.4	188	103	54.8 ± 4.9%
−1.4 to −1.2	301	184	61.1 ± 3.6%
−1.2 to −1	412	219	53.2 ± 3.4%
−1 to −0.8	655	278	42.4 ± 3.0%
−0.8 to −0.6	1074	339	31.6 ± 2.5%
−0.6 to −0.4	2205	535	24.3 ± 1.9%
−0.4 to −0.2	5492	499	9.1 ± 1.3%
−0.2–0	8254	378	4.6 ± 1.1%
0–0.2	11404	312	2.7 ± 0.9%
0.2–0.4	3541	72	2.0 ± 1.7%
0.4–0.6	260	8	3.1 ± 6.1%
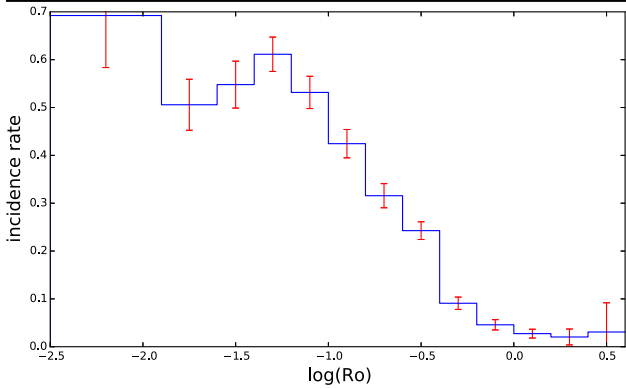
Notes. The left column shows the range of the logarithm of the Rossby number Ro, the second column shows the number of stars N_{stars} in the McQuillan et al. (2014b) database, the third column the number of flare stars N_{flare} (as detected in this work and cross-referenced with McQuillan et al. 2014b), and the fourth column is the flare star incidence rate per rotation bin. The graph shows the flare star incidence rate as a function of Rossby number Ro.

average 2.0 times per quarter, and flaring giants flare on average 1.6 times per quarter.

However, from the fit of the slope m to the histogram, it seems that all stellar types have a similar tail of flare stars that have a high flare occurrence rate. Thus, the colloquially used term "flare stars" seems to make sense with our statistics: each stellar type has a fraction of flare stars. Such flare stars are more frequently occurring in K- and M-type stars, but exist for earlier types as well. They are rarest among the giant class.

3.8. Relation with Rotation Period

Because Candelaresi et al. (2014) and Davenport (2016) found a strong correlation between the flare occurrence rate, luminosity, and rotation period, we also investigate the statistical connection between stellar flares and stellar rotation period. To that end, we have used the data of McQuillan et al. (2014a), who have found the stellar rotation from a period analysis of light curve variations due to stellar spots. We have cross-referenced our flare stars with the data available in McQuillan et al. (2014b). In total, we obtained the rotation period for 3033 of our flare stars from that catalog. We note that none of the flaring giants have reported rotation periods, because they were excluded from McQuillan et al. (2014a).

For each star, we estimate the Rossby number Ro by $Ro = p_{rot}/\tau$, where p_{rot} is the rotation period (taken from McQuillan et al. 2014b), and τ is the convective turnover time. The latter is estimated with formula 11 from Wright et al. (2011), following the procedure in Davenport (2016) but using the estimated stellar masses from the KIC.

Table 4 shows the number of flare stars for each bin and the flare star incidence rate. Here the standard deviation to the incidence rate p is estimated as $\sqrt{p(1-p)/N_{stars}}$. The incidence rates are also shown as a function of the Rossby number in the graph under Table 4. It is obvious that the stars with a short rotation period are much more often also flare stars, in comparison with stars with a long rotation period. However, for short rotation periods (low Rossby numbers), there is again a drop or saturation from the maximum at $\log Ro \sim -1.2$. This seems to compare well to the results of Candelaresi et al. (2014), where the maximum occurs around $\log Ro \sim -1$ (see their Figure 2). However, they plotted flare rate, instead of our occurrence rate. It also agrees with the saturation found by Davenport (2016).

From the graph in Table 4, it seems that the flare star incidence rate tails off to a constant value from a Rossby number of 1 onward, to a value of about 3%. An interesting question is, does the Sun fall in the 3% of active stars? Or would we not observe it as a flaring star with our detection algorithm? In any case, these reported incidence rates of 3% are an overestimate of the real occurrence rate, because the rotation period is measured by the appearance of stellar spots, thus subselecting a sample of only magnetically active stars.

In Figure 7 (top rows), we show the histograms for flare energy, flare duration, flare amplitude, and flare occurrence rate per bin of Rossby numbers (roughly corresponding to left of the incidence rate peak, the peak, two in the tail, and one for the slow rotators). The graphs show the distribution obtained using the least-squares fit or MLE/MLE$_1$ estimator (Equation (5)) as the olive line. The values of the MLE or slope of the least-squares fit are indicated in the key of the figure. In the bottom row, we show the values of MLE/MLE$_1$ as a function of the Rossby number.

The graph for the occurrence frequency (Figure 7, middle right panel) confirms the behavior of Table 4. If a star is rapidly rotating, it is not only highly likely to flare, but it is also more likely that it will flare more often than stars that are not fast rotators. Here as well, the saturation past $\log Ro \sim -1.2$ is apparent. In the middle left panel of Figure 7, we see that stellar rotation periods also have a strong influence on the flare amplitude. If a star is rotating quickly, it has a higher chance of creating a big flare.

Thus, the effect of rotation is twofold: quickly rotating stars have a higher chance that a flare is occurring (or that flares occur more often), and flares that occur have a higher chance of being strong flares.

On the other hand, the graphs for the flare energy (top left panel) show that the power-law slope does not depend on the Rossby number. Thus, this confirms the results of Candelaresi et al. (2014), who also found that the slope for the flare energy distribution is nearly constant. This was recently confirmed with LAMOST by Chang et al. (2017).

In the bottom panels of Figure 7, we show how the MLEs of the amplitude distribution and occurrence rate distribution vary with the rotation period. A very clear correlation is found for
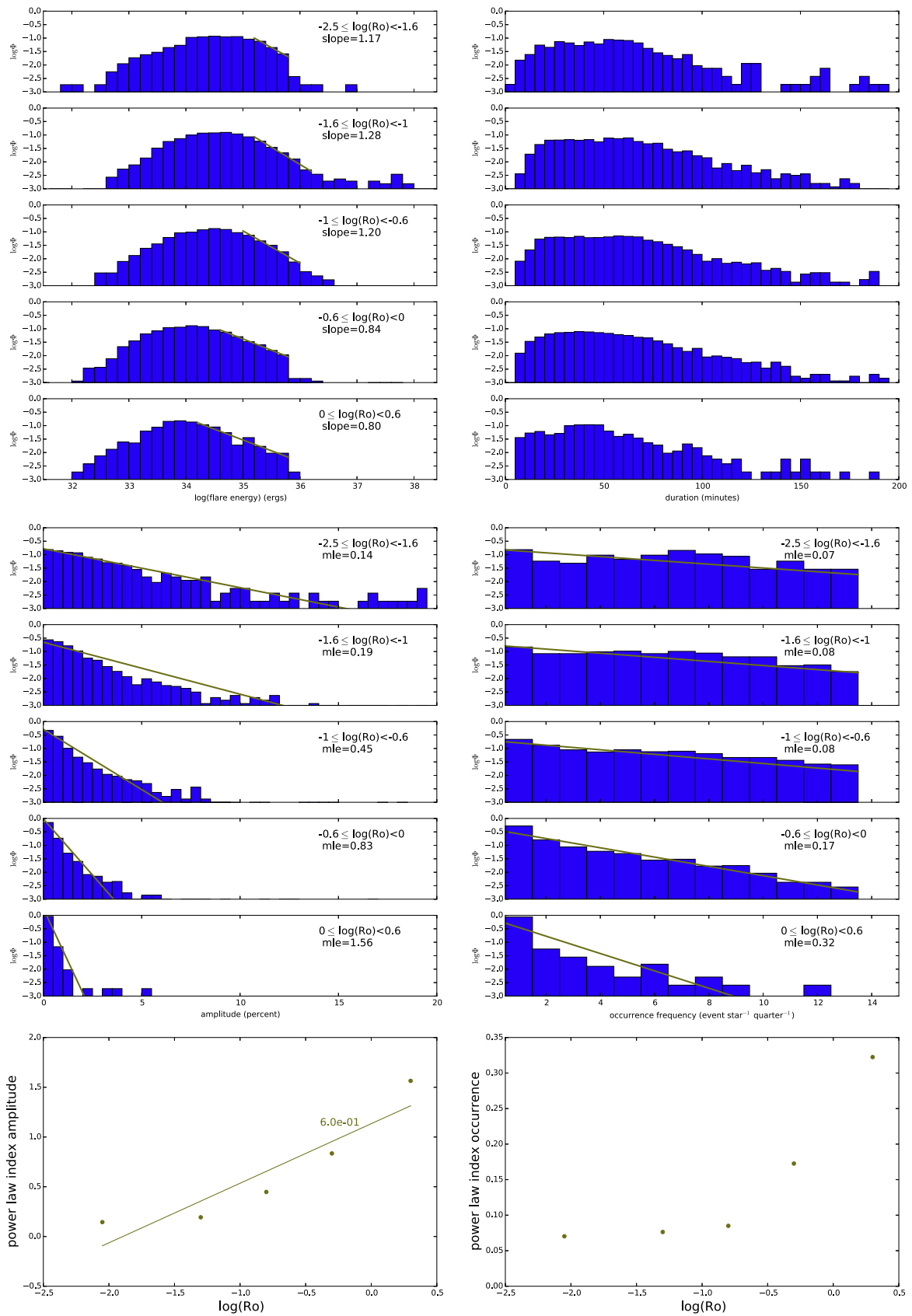
Figure 7. The top row shows the (logarithm of the) histogram of the flare energy (top left) and flare duration (top right), split up per bin in Rossby number. The second row shows the histogram of flare amplitude (middle left) and flare rate (middle right). The olive line shows the distribution given by the least-squares fit (top left), the MLE (middle left), and MLE$_1$ (middle right). The key contains the parameters of the least-squares fit and the MLE/MLE$_1$. The bottom row shows the MLE/MLE$_1$ estimator (olive dots) per Rossby bin, and the linear least-squares fit to those values in the bottom left panel. The indicated value is the slope of the linear least-squares fit.

the slower rotators, allowing us to accurately predict flare amplitude and flare occurrence rate distributions when a stellar rotation period is given. Perhaps such a scaling would modify the estimates for the occurrence of superflares on our Sun (Shibata et al. 2013; Notsu et al. 2015).

4. Conclusions

In this paper, we have developed a new method for automated flare detection, using thresholds in the intensity increase, the increase in the running difference, and the flare duration. We have applied it to the long-cadence data from quarter 15 of the *Kepler* mission, and we made the flare detections available to the wider community for further study in Table 1.

In Section 3, we have reported the discovery of several new flaring A stars, which had not been found by previous studies. From stellar evolution models, it is unexpected that A stars would flare, because they have a small or no surface convection zone, and thus a magnetic dynamo could not operate. Even though Pedersen et al. (2017) discarded some flare detections in A stars, our paper adds further to the evidence of Balona (2012), and these stars should be studied in detail in order to verify the stellar evolution theory.

Balona (2015) had found a few cases of flaring giants. We have extended the sample of flaring giants by 653 objects. It is unexpected that evolved stars would have strong magnetic fields, because they have spun down over their evolution. Moreover, the sheer size of these stars would also decrease the strength of magnetic fields generated in the core. Thus, observations of stellar flares on giants could be evidence that a surface dynamo is operating efficiently there. The giants show no rotational modulation, indicating that starspots do not cover a significant amount of the surface area.

We have then embarked on a statistical study of the flares in our sample. Regarding the flare amplitude, we have found that flares on F- and G-type stars behave similarly and are to be contrasted with K- and M-type stars (including giants). This could be an argument that the magnetic activity is just determined by the location in the HR diagram and that two causes for magnetic activity exist.

On the other hand, the flare duration could be grouped between F stars and giants on the one side and G, K, and M stars on the other side. This then seems to suggest that the flare activity is determined by the initial stellar mass and is kept while flare stars move from the main sequence to the giant branch. From the flare energy, we find the same power law for the G, K, and M dwarfs.

From the statistics of the flare occurrence rates, we have found that all spectral classes have a tail of stars that are frequently flaring. We could name these stars "flare stars." Such a tail of frequently flaring stars is more prominent for K and M types, but also easily found for earlier types. Flare stars are less frequently observed for giant stars.

We have cross-referenced our flaring stars with the rotation periods found in McQuillan et al. (2014b). By splitting the flare stars by Rossby number, it became very clear that the rotation period has a very large influence on stellar flares. A rapidly rotating star has a higher chance to flare and has a higher chance to flare more often, and the flares on rapidly rotating stars are often stronger than on slowly rotating stars. This matches earlier findings that the X-ray luminosity of

Table 5
New Incidence Rates Using Updated Stellar Parameters. The Table is Equivalent to Table 2

Stellar Type	# Objects	# Flare Stars	Incidence
A+B	2861	70	2.45%
F	51916	1230	2.37%
G	108463	3207	2.96%
K+M	36583	2146	5.87%
giants	21845	695	3.18%

stars is correlated with its rotation period (Wright et al. 2011), and thus flaring activity is also a good predictor of rotation period and X-ray luminosity. We confirm the findings of Candelaresi et al. (2014) and Davenport (2016) that there is a saturation of the dynamo for very rapidly rotating stars, or even a decrease in flaring activity for lower Rossby numbers. It remains to be seen how the chance for a superflare on our Sun is influenced when incorporating our statistical relation between flare amplitude and flare occurrence rate and the solar rotation period.

In the future, we want to extend our study to the full *Kepler* database. Such a larger sample and longer observation period will provide even better statistics than what we reported in this paper. This would enable us to split up the sample both in spectral classes and simultaneously with Rossby number, allowing for the disentangling of the effects of either on the stellar activity. Moreover, the full sample of the automated flare detection in *Kepler* should be cross-referenced with earlier works (such as Davenport 2016), which would make the flare database more robust.

To make the statistics of the sample more robust and disentangle binarity from the flaring nature, detailed spectroscopic studies of the flare stars are required to exclude the possibility of a companion (Karoff et al. 2016; Chang et al. 2017; Pedersen et al. 2017). In particular, the existence of flaring A stars conflicts with our current understanding of stellar evolution, and these could use special spectroscopic scrutiny.

Facility: Kepler.

Appendix A
Influence of KIC

In this appendix, we study the influence of stellar parameters taken from the KIC. To that end, we have taken the `q1_q17_dr25_stellar` catalog downloaded from https://exoplanetarchive.ipac.caltech.edu/, which contains the latest update to the stellar parameters.

With the new catalog, the incidence rates of flare stars over the different spectral types are slightly altered. Most incidence rates are not too far from the earlier reported values, except for the incidence rates for A and F stars. For the A stars, the
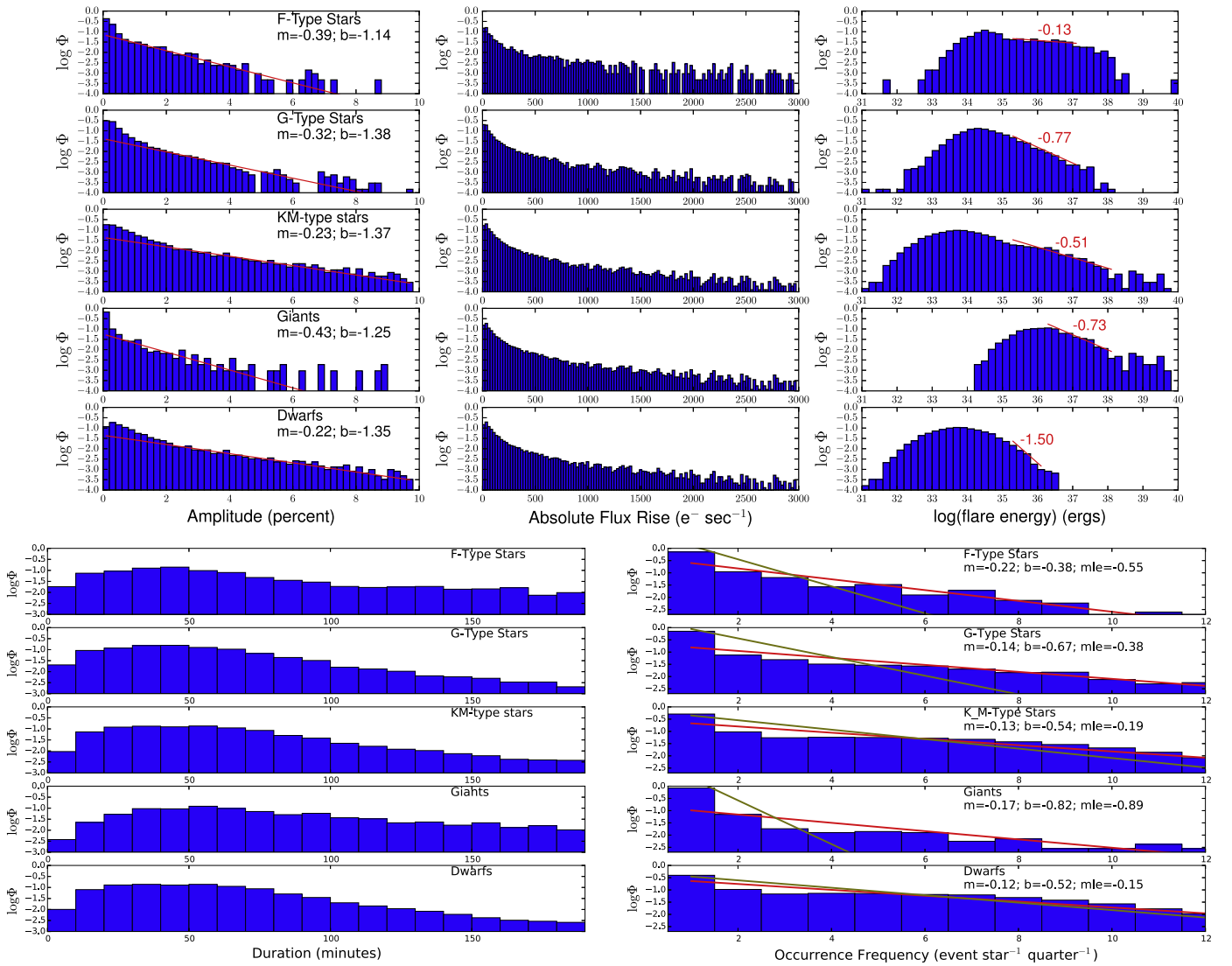
Figure 8. The equivalent figures using updated stellar parameters using the `q1_q17_dr25_stellar` catalog. The top row is the same as Figure 4, the bottom left is equivalent to Figure 5, and the bottom right is equivalent to Figure 6.

incidence rate seems to be nearly doubled, and the rate of the F stars is decreased. The reason is that several flare stars that previously had an F type have a slightly increased effective temperature in the new stellar parameter database. The new temperature crosses our (nearly arbitrarily chosen) threshold between F and A stars of 7600 K. In that sense, the numbers in Table 5 give a feeling of the sensitivity of the incidence rates to the choice of temperature threshold between spectral types and perturbations in the catalog temperatures. Probably a slightly adjusted threshold temperature would return the incidence rates to the earlier values.

We have also used the updated stellar parameter catalog to study the flare parameters. The updated versions of Figures 4–6 are displayed in Figure 8. From the comparison of the updated figures with Figure 4, it seems that the power-law slope of the amplitude distribution is not changed, except for the giants. However, the slopes of the flare energy distribution are very sensitive to the catalog parameters, and therefore they are probably not so robust. The flare duration appears to be insensitive to the update, and so does the occurrence frequency.

Appendix B
Flare Light Curves for A Stars

In this appendix, we show the flare light curves for all detected flares on A stars, together with their pixel mask. In each of the figures, we show the intensity time series in the left panel. The time of the flare is indicated with a vertical dotted line. The middle panel shows the pixel data for the flare. The right panel shows the pixel data for the star. The latter is constructed as the average of the pixel intensities between two and five data points before the flare peak time. The data in the middle panel is the difference of the pixel intensity at the flare peak time minus the mean pixel intensity shown in the right panel. In both the middle and right panels, the intensity is masked to only show intensity that was added to obtain the intensity time series in the left panel.

Some of the light curves of the A stars show periodic modulation of the light curve (see, e.g., Figure 9.15 or 9.25). However, it is unclear whether these variations are due to rotational modulation due to stellar spots or periodic modulation due to stellar pulsations.
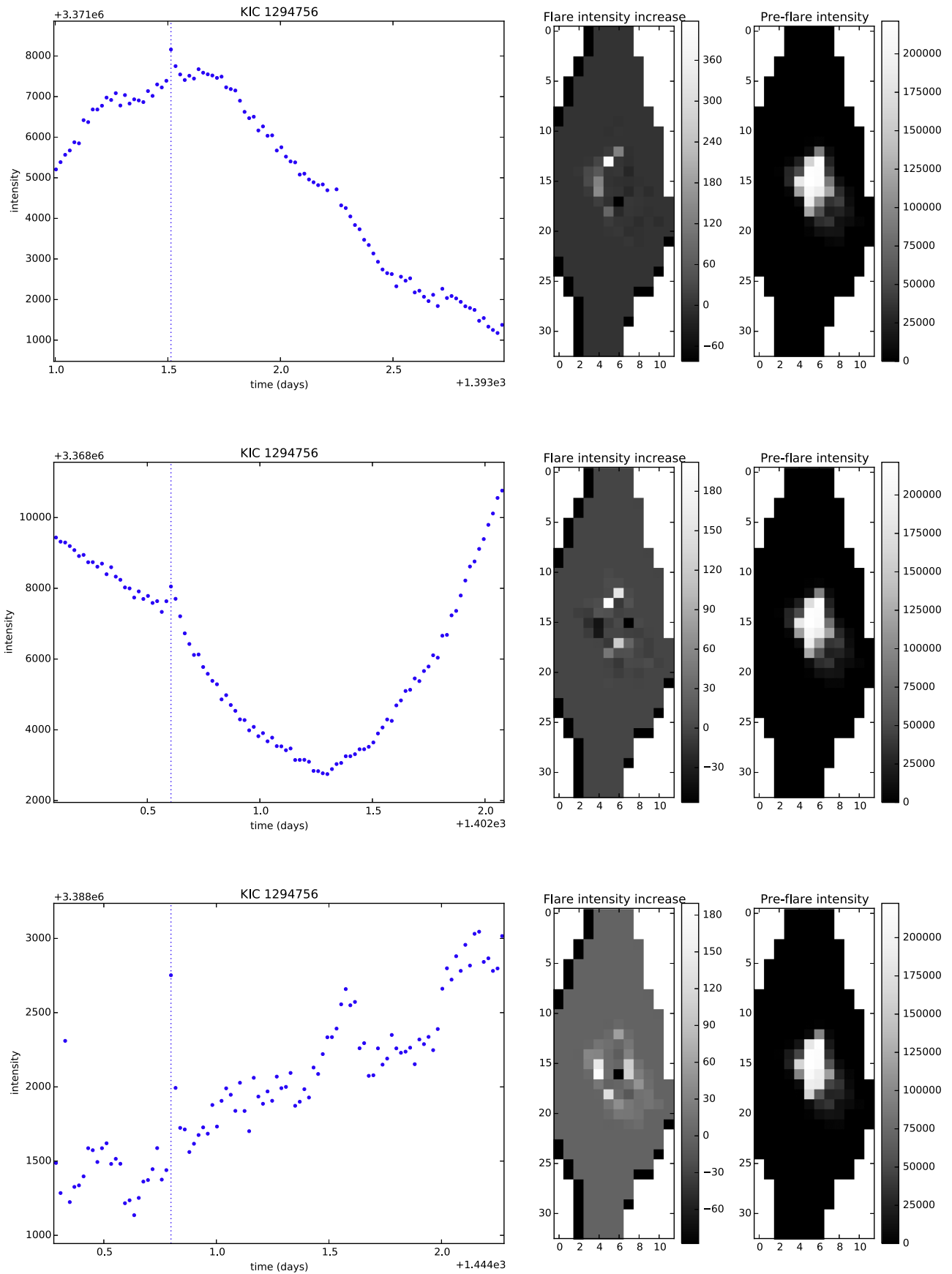
Figure 9. Flare light curves for KIC 1294756.

(The complete figure set (28 images) is available.)

ORCID iDs

Tom Van Doorsselaere ◉ https://orcid.org/0000-0001-9628-4113

References

Amari, T., Luciani, J.-F., & Aly, J.-J. 2015, Natur, 522, 188
Anfinogentov, S., Nakariakov, V. M., Mathioudakis, M., Van Doorsselaere, T., & Kowalski, A. F. 2013, ApJ, 773, 156
Balona, L. A. 2012, MNRAS, 423, 3420
Balona, L. A. 2013, MNRAS, 431, 2240
Balona, L. A. 2015, MNRAS, 447, 2714
Balona, L. A., Broomhall, A.-M., Kosovichev, A., et al. 2015, MNRAS, 450, 956
Candelaresi, S., Hillier, A., Maehara, H., Brandenburg, A., & Shibata, K. 2014, ApJ, 792, 67
Chang, H.-Y., Song, Y.-H., Luo, A.-L., et al. 2017, ApJ, 834, 92
Cho, I.-H., Cho, K.-S., Nakariakov, V. M., Kim, S., & Kumar, P. 2016, ApJ, 830, 110
Davenport, J. R. A. 2016, ApJ, 829, 23
Degroote, P., Briquet, M., Catala, C., et al. 2009, A&A, 506, 111
Gaulme, P., Jackiewicz, J., Appourchaux, T., & Mosser, B. 2014, ApJ, 785, 5
Harper, G. M., O'Riain, N., & Ayres, T. R. 2013, MNRAS, 428, 2064
Hawley, S. L., Davenport, J. R. A., Kowalski, A. F., et al. 2014, ApJ, 797, 121
Johnson, N. L., & Kotz, S. 1970, Continuous Univariate Distributions (New York: Wiley)
Karoff, C., Knudsen, M. F., De Cat, P., et al. 2016, NatCo, 7, 11058
Karovska, M., Schlegel, E., Hack, W., Raymond, J. C., & Wood, B. E. 2005, ApJL, 623, L137
Konstantinova-Antova, R., Aurière, M., Iliev, I. K., et al. 2008, A&A, 480, 475

Kowalski, A. F., Hawley, S. L., Holtzman, J. A., Wisniewski, J. P., & Hilton, E. J. 2010, ApJL, 714, L98
Kowalski, A. F., Hawley, S. L., Wisniewski, J. P., et al. 2013, ApJS, 207, 15
Lurie, J. C., Davenport, J. R. A., Hawley, S. L., et al. 2015, ApJ, 800, 95
Maehara, H., Shibayama, T., Notsu, Y., et al. 2015, EP&S, 67, 59
Maehara, M., Shibayama, T., Notsu, S., et al. 2012, Natur, 485, 478
Mathioudakis, M., Seiradakis, J. H., Williams, D. R., et al. 2003, A&A, 403, 1101
McQuillan, A., Mazeh, T., & Aigrain, S. 2014a, ApJS, 211, 24
McQuillan, A., Mazeh, T., & Aigrain, S. 2014b, yCat, 221, 10024
Mitra-Kraev, U., Harra, L. K., Williams, D. R., & Kraev, E. 2005, A&A, 436, 1041
Notsu, Y., Honda, S., Maehara, H., et al. 2015, PASJ, 67, 33
Pedersen, M. G., Antoci, V., Korhonen, H., et al. 2017, MNRAS, 466, 3060
Pitkin, M., Williams, D., Fletcher, L., & Grant, S. D. T. 2014, MNRAS, 445, 2268
Pugh, C. E., Armstrong, D. J., Nakariakov, V. M., & Broomhall, A.-M. 2016, MNRAS, 459, 3659
Pugh, C. E., Nakariakov, V. M., & Broomhall, A.-M. 2015, ApJL, 813, L5
Ramsay, G., & Doyle, J. G. 2014, MNRAS, 442, 2926
Ramsay, G., Doyle, J. G., Hakala, P., et al. 2013, MNRAS, 434, 2451
Shakhovskaia, N. I. 1989, SoPh, 121, 375
Shibata, K., Isobe, H., Hillier, A., et al. 2013, PASJ, 65, 49
Shibayama, T., Maehara, H., Notsu, S., et al. 2013, ApJS, 209, 5
Simon, T., & Drake, S. A. 1989, ApJ, 346, 303
Srivastava, A. K., Lalitha, S., & Pandey, J. C. 2013, ApJL, 778, L28
Sun, J. Q., Cheng, X., Ding, M. D., et al. 2015, NatCo, 6, 7598
Van Doorsselaere, T., Kupriyanova, E. G., & Yuan, D. 2016, SoPh, 291, 3143
Vlemmings, W. H. T., Ramstedt, S., O'Gorman, E., et al. 2015, A&A, 577, L4
Walkowicz, L. M., Basri, G., Batalha, N., et al. 2011, AJ, 141, 50
Welsh, B. Y., Wheatley, J., Browne, S. E., et al. 2006, A&A, 458, 921
Wright, N. J., Drake, J. J., Mamajek, E. E., & Henry, G. W. 2011, ApJ, 743, 48