Draft version August 12, 2024
Typeset using LaTeX default style in AASTeX631

arXiv:2408.05038v1 [astro-ph.SR] 9 Aug 2024

Detecting quasi-periodic pulsations in solar and stellar flares with a neural network

Sergey A. Belov [1], Dmitrii Y. Kolotkov [1,2], Valery M. Nakariakov [1,2] and Anne-Marie Broomhall [1]

[1] Centre for Fusion, Space and Astrophysics, Department of Physics, University of Warwick, Coventry CV4 7AL, UK
[2] Engineering Research Institute "Ventspils International Radio Astronomy Centre (VIRAC) ", Ventspils University of Applied Sciences, Ventspils, LV-3601, Latvia

ABSTRACT

Quasi-periodic pulsations (QPP) are often detected in solar and stellar flare lightcurves. These events may contain valuable information about the underlying fundamental plasma dynamics as they are not described by the standard flare model. The detection of QPP signals in flare lightcurves is hindered by their intrinsically non-stationary nature, contamination by noise, and the continuously increasing amount of flare observations. Hence, the creation of automated techniques for QPP detection is imperative. We implemented the Fully Convolution Network (FCN) architecture to classify the flare lightcurves whether they have exponentially decaying harmonic QPP or not. To train the FCN, 90,000 synthetic flare lightcurves with and without QPP were generated. After training, it showed an accuracy of 87.2% on the synthetic test data and did not experience overfitting. To test the FCN performance on real data, we used the subset of stellar flare lightcurves observed by Kepler, with strong evidence of decaying QPP identified hitherto with other methods. Then, the FCN was applied to find QPPs in a larger-scale Kepler flare catalogue comprised of 2274 events, resulting in a 7% QPP detection rate with a probability above 95%. The FCN, implemented in Python, is accessible through a browser application with a user-friendly graphical interface and detailed installation and usage guide. The obtained results demonstrate that the developed FCN performs well and successfully detects exponentially decaying harmonic QPP in real flare data, and can be used as a tool for preliminary sifting of the QPP events of this type in future large-scale observational surveys.

Keywords: Sun: flares — Sun: oscillations — stars: flare — techniques: miscellaneous

1. INTRODUCTION

Solar flares, together with coronal mass ejections, are the most powerful physical processes in the solar system (e.g., Benz 2017). Understanding the physical processes which cause the impulsive releases of the magnetic energy in the solar atmosphere are among the key challenges of modern solar physics. Furthermore, there is a growing interest in similar phenomena detected on other stars, including the sun-like stars with potentially habitable planetary systems. The so-called standard model of a solar flare attributes the energy release to the process of magnetic reconnection (e.g., Shibata & Magara 2011).

An intensively studied phenomenon which is not described by the standard flare model are quasi-periodic pulsations (QPP) of the flaring emission (e.g., Nakariakov & Melnikov 2009; Nakariakov et al. 2010; Zimovets et al. 2021). QPP appear as the quasi-periodic modulation of the flaring emission in all observational bands, from radio to gamma-rays, in both thermal and non-thermal emission, and in flares of all classes, from microflares (e.g., Nakariakov et al. 2018) to the most-power flare classes (e.g., Kolotkov et al. 2018). Typical periods of QPP range from a fraction of a second to several minutes. Simões et al. (2015) demonstrated that about 80% of X-class solar flares display QPP in the soft X-ray emission. Furthermore, Hayes et al. (2020) found out that approximately 46% of X-class, 29% of M-class, and 7% of C-class flares show evidence of stationary, i.e., narrowband QPP. Those estimations are highly sensitive to the

Corresponding author: Sergey A. Belov
Sergey.Belov@warwick.ac.uk

detection criterion, i.e., upon the definition of a QPP. One of the complications is the essentially non-stationary nature of QPP (Nakariakov et al. 2019a; Pascoe et al. 2020; Mehta et al. 2023). The QPP patterns may have pronounced amplitude and/or instantaneous oscillation period modulation, and also the signal shape may be highly anharmonic. It means that the oscillation energy is intrinsically distributed over several Fourier frequencies, making the detection techniques based on the Fourier transform not very relevant (e.g., Anfinogentov et al. 2022).

Zimovets et al. (2021) identified at least fifteen mechanisms which could be responsible for the appearance of QPP. These mechanisms include the modulation of the magnetic reconnection rate or parameters of the emitting plasma by magnetohydrodynamic (MHD) waves, and self-induced repetitive reconnection (McLaughlin et al. 2018). As different mechanisms produce QPP with different properties, such as oscillation periods, signal shapes, modulation, etc., one would expect the existence of several different classes of QPP. However, the task of the QPP taxonomy remains an outstanding problem. Obviously, the detection technique must be fine-tuned for the detection of QPP of a specific class. Similarly, the search for statistical relationships among different observables should be conducted within a specific QPP class.

One of the clearly identified class are QPP of co-called SUMER oscillations, named after the instrument used in the first detection of this oscillatory phenomenon (e.g., Wang 2011; Wang et al. 2021). In spectral observations, SUMER oscillations appear as periodic alternate Doppler shift of a coronal emission line. The oscillation pattern is highly harmonic. A typical period of the SUMER oscillations is about several minutes. The oscillation decays very rapidly, with the exponential damping time being comparable to the oscillation period. SUMER oscillations are interpreted as standing slow magnetoacoustic waves in a coronal loop. A related phenomenon are sloshing oscillations, characterised by a slow magnetoacoustic wave bouncing between footpoints of the loop (e.g., Reale et al. 2019). QPP of the SUMER class appear in the decay phase of a flare as, e.g., the modulation of the intensity of the soft X-ray radiation, EUV radiation associated with hot plasma, and radio (e.g., Kim et al. 2012). Similar QPP are detected in the soft X-ray radiation produced by stellar flares. The linear scaling of the damping time with the oscillation period has been established in both solar and stellar flares (Cho et al. 2016). Furthermore, similar QPP are detected in the white light emission produced by stellar flares (e.g., Pugh et al. 2016; Bai et al. 2023).

The search for QPP in solar flares is complicated by a large number of events, which is typically about 10,000 of C-, M-, and X-class flares during an eleven-year cycle, and much more lower-power flares. On the other hand, the large number of events allows for the statistical validation and comparison of various theories, and also creates a ground for the application of machine learning (ML) techniques. ML techniques have been intensively applied to various tasks of modern solar physics for several years. It includes applications to space weather forecasting, including solar flares (e.g., Ahmed et al. 2013; Bobra & Couvidat 2015; Nishizuka et al. 2018; Camporeale 2019) and coronal mass ejections (e.g., Wang et al. 2019; Georgoulis et al. 2021), and predictors of the geomagnetic activities (e.g., Valach et al. 2009); identification of morphological features in prominences, and their relationship with solar activity (e.g., Zhang et al. 2024); image super-resolution techniques (e.g., Xu et al. 2024); recognition of coronal loops (e.g., Wang et al. 2024); inferring transverse velocities at the solar surface (e.g., Tilipman et al. 2023); extrapolation of solar global magnetic fields(e.g., Jeong et al. 2020); and may others. Machine learning techniques have also been applied to the detection of stellar flares in photometric surveys (e.g., Vida & Roettenbacher 2018), see also Fluke & Jacobs (2020), and the search for quasi-periodic eruptions in the X-ray emission from the nuclei of galaxies (e.g., Webbe & Young 2023). Thus, the application of an ML approach may significantly advance the QPP study too. Promising research avenues are, for example, an ML-based identification of QPP classes, and the application of a pattern-recognition approach to the QPP detection.

The aim of this paper is to develop and test an ML technique for the detection of QPP patterns of the SUMER class, i.e., rapidly-decaying highly harmonic oscillatory patterns in lightcurves, and provide the research community with an open access software package implementing this technique. The paper is organised as follows. In Section 2 we describe the synthetic data which are used for training a neural network. Section 3 describes the neural network architecture together with its performance on the synthetic data. In Section 4, we apply the developed neural network to real flare lightcurves. Then, Section 5 contains the information about the browser application to use the network. Finally, Section 6 concludes the results and outlines the possible next steps to develop the approach presented in this paper.

2. SYNTHETIC DATA CREATION

Usually, a large amount of data is required for training and validating deep learning (DL) models. For our problem, the dataset should contain at least several thousands of lightcurves with and without QPP. Moreover, in training, we need to provide the network with a so-called ground truth knowledge, i.e., a definitive answer as to whether the given flare lightcurve contains a QPP signal or not, which is not always available in the existing flare catalogues. For this reason, we construct an extensive synthetic dataset of flare lightcurves and rely on it for training and validating our network. To generate the synthetic dataset, we follow the procedure similar to that presented in Broomhall et al. (2019), which is briefly outlined below.

Our synthetic flare lightcurves have the following general form,

$$\mathcal{I}\left(t\right) = \mathcal{F}\left(t\right) + \mathcal{Q}\left(t\right) + \mathcal{N}\left(t\right). \tag{1}$$

Here, $\mathcal{F}(t)$ is the flare profile with the peak amplitude $A_{\rm flare}$, generated using one of the three flare shape models proposed hitherto,

$$\mathcal{F}_1(t) = A_{\rm flare}\frac{\mathrm{G}\left(t\right)}{\mathrm{Max}\left(\mathrm{G}\left(t\right)\right)}, \tag{2}$$
$$\mathrm{G}\left(t\right) = \exp\left[D\left(B-t\right)+\frac{C^2D^2}{4}\right]\left[\mathrm{erf}\left(Z\right)-\mathrm{erf}\left(Z-\frac{t}{C}\right)\right],$$
$$Z = \left(2B+C^2D\right)/2C,$$

based on the convolution of the Gaussian energy deposition and exponential energy dissipation with B, C, D being free parameters (Gryciuk et al. 2017),

$$\mathcal{F}_2(t) = \begin{cases} A_{\rm flare}\exp\left(-\dfrac{t^2}{2\sigma_{\rm rise}}\right), & t<0 \\ A_{\rm flare}\exp\left(-\dfrac{t^2}{2\sigma_{\rm decay}}\right), & t\geq 0 \end{cases} \tag{3}$$

which is a guessed flare profile approximated by two half-Gaussian curves of different widths $\sigma_{\rm rise}$ and $\sigma_{\rm decay}$ (Broomhall et al. 2019), and

$$\mathcal{F}_3(t) = \begin{cases} A_{\rm flare}\left(1+1.941t-0.175t^2-2.246t^3-1.125t^4\right), & -1\leq t<0 \\ 0.948A_{\rm flare}\exp\left(-0.965t\right), & 0\leq t<1.6 \\ 0.322A_{\rm flare}\exp\left(-0.290t\right), & 1.6\leq t\leq 19 \end{cases} \tag{4}$$

derived as an empirical template of a stellar flare in Kepler observations (Davenport et al. 2014). Then, the flare duration is scaled from its initial value of 20 to the new length $L_{\rm flare}$ and the flare peak is shifted to the right by a value determined by the "shift" parameter (see Table 1). As a final step, the constant offset value is added to the flare.

For the QPP signal $\mathcal{Q}(t)$ in Eq. (1), in this work we focus on a particular type of rapidly decaying harmonic QPP events in the flare decay phase, proposed by Nakariakov et al. (2019a) and which can be associated with the dynamics of standing slow magnetoacoustic waves, also known as SUMER oscillations, in a flaring loop (e.g., Wang et al. 2021; Nakariakov et al. 2019b),

$$\mathcal{Q}\left(t\right) = A_{\rm qpp}\exp\left(-\frac{t}{\tau}\right)\cos\left(\frac{2\pi t}{P}+\phi\right). \tag{5}$$

In Eq. (5), $A_{\rm qpp}$ is the QPP amplitude; P, τ, and ϕ stand for the QPP period, decay time, and phase, respectively. The QPP start time $t_{\rm start}$ is chosen to be about the time of the flare peak $t_{\rm peak}$, but not exactly coinciding with it (see the parameter $t_{\rm start}-t_{\rm peak}$ in Table 1). Next, the noise component $\mathcal{N}(t)$ is introduced into the lightcurve, which is modelled as either white noise, red noise, or a combination of both types of noise. The red noise component is generated as

$$\mathcal{N}_i = r\mathcal{N}_{i-1}+\sqrt{(1-r^2)}w_i, \tag{6}$$

where $\mathcal{N}_i$ is the red noise value at the i-th instant of time, obtained from a white-noise component w_i, with the correlation coefficient between successive data points, r. The white-noise component w_i is generated from a Gaussian distribution with zero mean and a standard deviation scaled relatively to the flare amplitude.

Table 1. Parameters of synthetic dataset.

Parameters	Gryciuk $\mathcal{F}_1(t)$	Two half-Gaussians $\mathcal{F}_2(t)$	Davenport $\mathcal{F}_3(t)$
$L_{\rm flare}$	$U\,(50, 300)$	$U\,(50, 300)$	$U\,(50, 300)$
shift	$U\,(0, 300 - L_{\rm flare})$	$U\,(0, 300 - L_{\rm flare})$	$U\,(0, 300 - L_{\rm flare})$
$A_{\rm flare}$	$10 + N\,(0, 2)$	$10 + N\,(0, 2)$	$10 + N\,(0, 2)$
$\sigma_{\rm rise}$	n/a	$U\,(1, 3)$	n/a
$\sigma_{\rm decay}$	n/a	$U\,(5, 20)$	n/a
B	$U\,(0, 5)$	n/a	n/a
C	$U\,(1, 5)$	n/a	n/a
D	$U\,(0.5, 1)$	n/a	n/a
offset	$U\,(0, 100)$	$U\,(0, 100)$	$U\,(0, 100)$
$A_{\rm qpp}/A_{\rm flare}$	$U\,(0.05, 0.5)$	$U\,(0.05, 0.5)$	$U\,(0.05, 0.5)$
$L_{\rm flare}/P$	$U\,(5, 15)$	$U\,(5, 15)$	$U\,(5, 15)$
τ/P	$U\,(1, 5)$	$U\,(1, 5)$	$U\,(1, 5)$
ϕ	$U\,(0, 2\pi)$	$U\,(0, 2\pi)$	$U\,(0, 2\pi)$
$t_{\rm start} - t_{\rm peak}$	$U\,(-0.05L_{\rm flare}, 0.1L_{\rm flare})$	$U\,(-0.05L_{\rm flare}, 0.1L_{\rm flare})$	$U\,(-0.05L_{\rm flare}, 0.1L_{\rm flare})$
White S/N	$U\,(1, 5)$	$U\,(1, 5)$	$U\,(1, 5)$
Red S/N	$U\,(1, 5)$	$U\,(1, 5)$	$U\,(1, 5)$
r	$U\,(0.81, 0.99)$	$U\,(0.81, 0.99)$	$U\,(0.81, 0.99)$

To train the ML model, we created 90,000 synthetic flare lightcurves using Eqs. (1)–(6), in which 50% of lighcurves have QPP and the other 50% of lighcurves do not (the QPP amplitude $A_{\rm qpp}$ set to zero). Each synthetic flare lightcurve contains 300 data points. In our dataset, we use nine combinations of the above-mentioned flare shapes $\mathcal{F}_{1,2,3}(t)$ and white/red/white+red noise $\mathcal{N}(t)$. The synthetic flare, QPP and noise parameters we used for the creation of our dataset are summarised in Table 1, where $U\,(a, b)$ stands for a random value uniformly distributed between a and b, and $N\,(\mu, \sigma)$ denotes a random value normally distributed with a mean μ and a standard deviation σ. The examples of the flare lightcurves with QPP, generated as described above, are shown in Fig. 1, for all three flare profile models $\mathcal{F}_{1,2,3}(t)$. The left panel of Fig. 2 shows 2400 randomly selected synthetic flare lightcurves plotted together, for illustration of the general shape of our dataset. To create this plot, the peaks of the selected lightcurves were positioned at one-third of the time domain and then scaled to unity.

To finalise our synthetic dataset, the created flare signals were randomly shuffled and split into train, test and validation subsets in a proportion of 80%, 10% and 10%, respectively. The synthetic dataset can be accessed online via Harvard Dataverse repository[1], where one can find all necessary source and data files and description.

3. NEURAL NETWORK ARCHITECTURE AND PERFORMANCE

In this work, we use the Fully Convolutional Network (FCN) architecture proposed by Wang et al. (2017) for a time series classification task. Our PyTorch Lightning implementation is based on the TensorFlow implementation from Ismail Fawaz et al. (2019). The FCN architecture is shown in Fig. 3 and consists of 3 consequent blocks of 1D convolution, batch normalization and Rectified Linear Unit (ReLu) activation function followed by 1D average pooling, one fully connected layer and a sigmoid activation function. As a result, an input times series is transformed into a QPP probability estimation $\mathcal{P}_{\rm QPP}$ ranging from 0 to 1. It is usually assumed that if this probability exceeds 0.5 than the ML model finds a positive class.

To increase the FCN performance and capability to learn from the data, we added one additional channel to the input data. This channel is produced by detrending the original flare time series. The algorithm of making this channel is as follows:

1. A position of a flare peak is found;

[1] doi.org/10.7910/DVN/UNRTN6

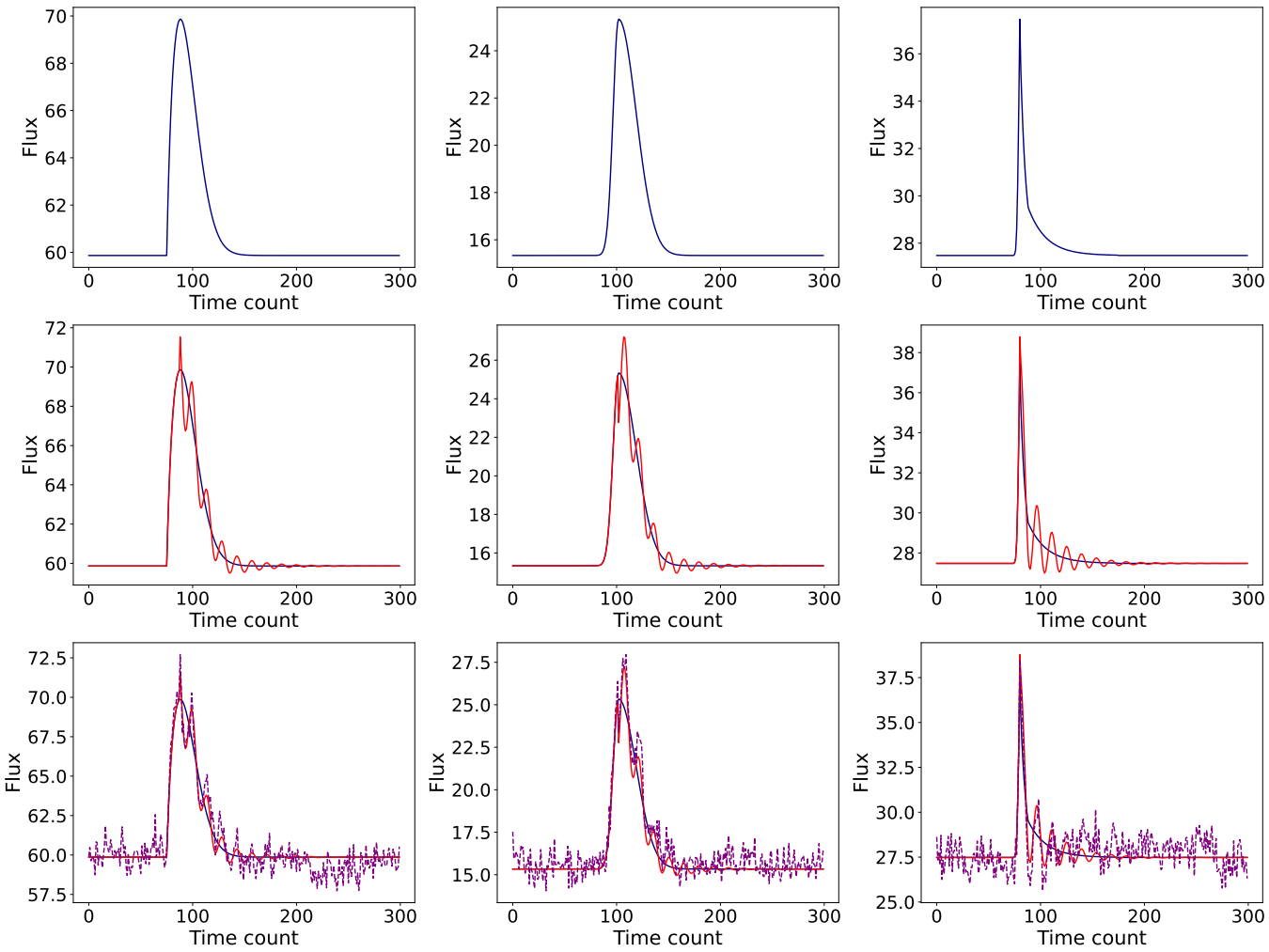

Figure 1. Examples of synthetic lightcurves given by flare models Eq. (2), left, Eq. (3), middle, and Eq. (4), right. The top row: the initial flare profiles. The middle row: the initial flare profiles with added QPP (red curve), see Eq. (5). The bottom row: the same as in middle row but with added red and white noise (purple dashed curve), see Eq. (6).

2. The flare is separated into the rise phase (before the flare peak) and decay phase (after the flare peak);

3. The decay and rise phase signals are smoothed by the Savitzky–Golay filter of the 4th order and a window width equal to the half-length of the input signal (separately for the rise and decay signals) to produce corresponding trends. For the illustration of smoothing, see the blue dash-dotted (decay phase) and black dashed (rise phase) curves in the left panel of Fig. 4;

4. For the decay phase, the obtained trend is subtracted from the original signal to enhance the visibility of QPP (see the blue curve on the right panel of Fig. 4);

5. For the rise phase, the detrended signal is also obtained, and its standard deviation σ is calculated. Then, the whole rise phase is padded with white noise produced with $U(-\sigma, \sigma)$ (see the black curve on the right panel of Fig. 4).

We found that adding the second channel improves the accuracy of the classification by about 5% on our synthetic dataset by highlighting the QPP features hidden by the trend (i.e. small amplitudes, long periodicities, etc.). It should be mentioned that, in general, this channel is sensitive to the choice of the Savitzky–Golay filter's order and window width (see e.g. the dedicated discussion of this issue in Sec. 4.2 of Kupriyanova et al. 2020). Our choice of these parameters allows for obtaining a smooth trend following slow (in comparison with QPP) changes in the

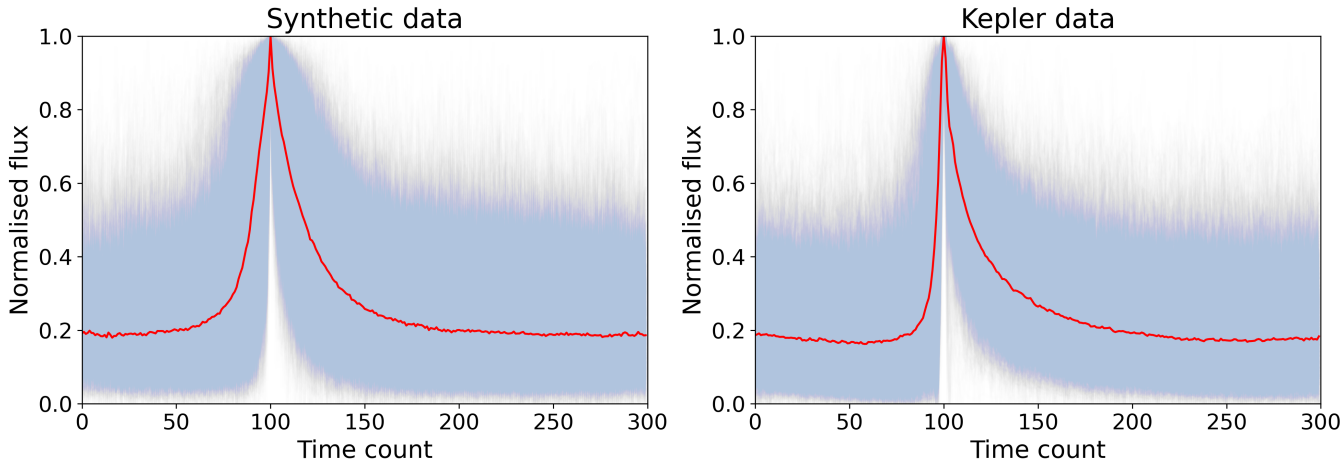

Figure 2. Simultaneous plot of 2400 aligned and min-max scaled light curves from the synthetic dataset (left panel) and min-max scaled light curves from stellar flare catalogue, observed with Kepler (Balona 2015).

lightcurve. Moreover, every window filter suffers from the edge effects which were mitigated by interpolating the edges by a polynomial function (the default 'interp' mode in the 'savgol_filter' function in the SciPy Python package). Nevertheless, the detrended time series is used as an additional complementary channel (in addition to the original time series) in this work, allowing the network to catch the time series properties which may be less pronounced in the original signal. The latter, complementary use of the detrended signal lets us ease the restrictions of the detrending and subjectivity in the choice of the filter's parameters.

Thus, the input data for the FCN consists of two channels. The first channel is the original flare lightcurve (e.g., the left panel of Fig. 4), and the second channel is the flare lightcurve detrended as described above (e.g., the right panel of Fig. 4). Before feeding into the network, channels in a data sample are standardised independently:

$$\hat{\mathcal{I}}^j_{c,i} = \left(\mathcal{I}^j_{c,i} - M\mathcal{I}^j_c\right)/\sigma\mathcal{I}^j_c, \tag{7}$$

where $\mathcal{I}^j_{c,i}$ is the c-th data channel of the j-th data sample at the i-th time count, and $M\mathcal{I}^j_c$ and $\sigma\mathcal{I}^j_c$ are mean and standard deviation of the time series $\mathcal{I}^j_{c,i}$.

The FCN has been trained with 64 sample batches on the train dataset using Adam optimizer (Kingma & Ba 2014) with learning rate of 0.001. The training progress has been tracked using the validation dataset. The early stopping has been used to stop the training process in the validation loss minimum. Figure 5 shows the dynamics of the loss and accuracy score calculated for the train and validation datasets during the training. It can be seen that the train and validation curves are close to each other, indicating that the FCN has generalized well to new (unseen) data and has not overfitted. Then, the FCN performance has been estimated using the test dataset. The test confusion matrices for different threshold values are shown in Fig. 6. For the 0.5 threshold, the FCN accuracy is 87.2% while precision is 93.6% (the ratio between the number of true positive answers and a number of all positive answers given by the FCN). For higher thresholds, accuracy decreases (but remains acceptable) with a simultaneous decrease in the number of false positives making the FCN more conservative.

Thus, we can conclude that the FCN performs well on the synthetic dataset and does not experience overfitting after the training. It indicates the capability of the FCN to generalise the data unseen during the training and makes it potentially applicable to the lightcurves where similar data patterns exist.

4. REAL DATA EXAMPLES

In this section, we demonstrate the application of our FCN to the existing real data QPP catalogues, observed in solar flares (the Automated Flare Inference of Oscillations (AFINO) catalogue[2], Inglis et al. 2015, 2016, using GOES soft X-ray data) and in stellar flares (Balona 2015; Pugh et al. 2016, with Kepler in white light).

[2] aringlis.github.io/AFINO

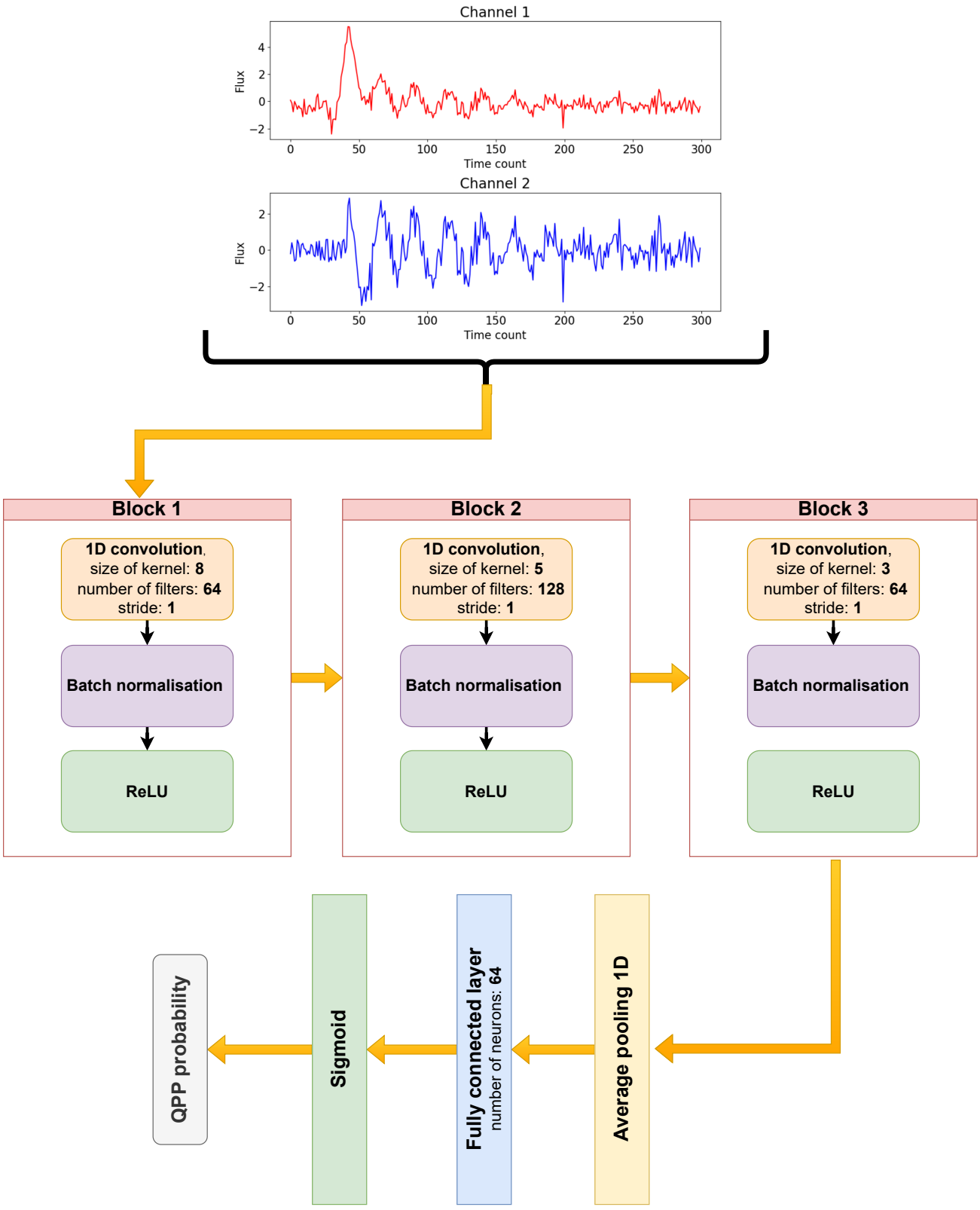

Figure 3. Architecture of the FCN for a QPP detection task.

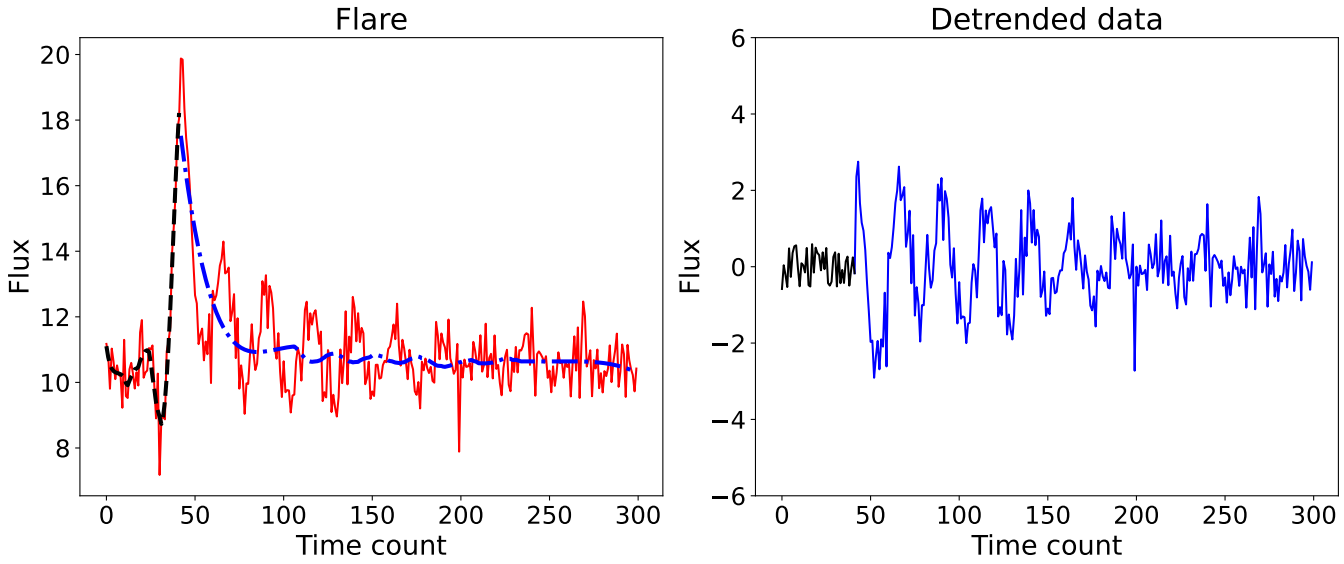

Figure 4. Preprocessing of a flare lightcurve to produce two input channels for our FCN (see the architecture in Fig. 3). Left panel: an initial flare time series (red curve) as the input channel 1, with the rise (black dashed curve) and decay (blue dash-dotted curve) trends. Right panel: the detrended flare lightcurve as the input channel 2, consisting of the white-noise padded rise phase (black curve) and the decay phase detrended with the Savitzky—Golay filter (blue curve).

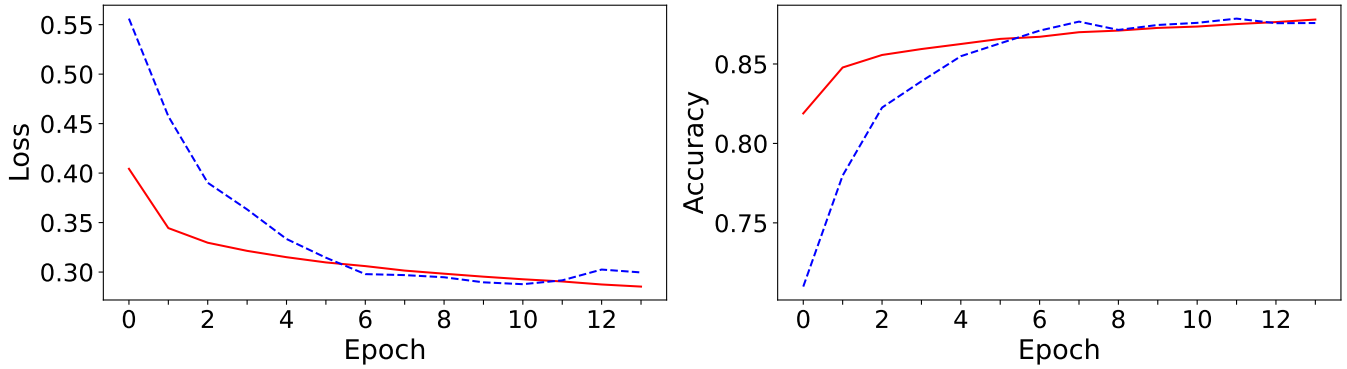

Figure 5. Left panel: the dynamics of the loss function during the training process for the train (red curve) and validation (blue dashed curve) datasets. Right panel: the dynamics of the accuracy score on the train (red curve) and validation (blue dashed curve) datasets during the training process.

4.1. *AFINO data*

To create the AFINO catalogue, the Fourier Power Spectral Density (PSD) was calculated for each observed flare time series. Then, the obtained PSD was fitted by three PSD models independently. These models are a single power-law plus a constant (model S0), a single power-law plus a constant and a localized enhancement associated with a QPP signature (model S1), and a broken-power law plus a constant (model S2). For each model, the Bayesian Information Criterion was calculated and the most significant model was chosen.

In our study, we target the rapidly decaying QPP events with periods longer than one minute (motivated by typical observed behaviour of QPP associated with SUMER oscillations in solar and stellar flares, see e.g., Cho et al. 2016). In contrast, in the AFINO catalogue, the mean period of QPP events is about 38 s, and only 79 out of 854 events have periods longer than one minute. Moreover, the QPP events in the AFINO catalogue are sufficiently long-lived, which is required by the PSD-based detection method. Thus, to look at the most relevant AFINO data, we chose 30 examples from this sub-set of 79 longer-period QPP events, 15 of which have strong preference for model S1 over models S0 and S2 (i.e., QPP detection with high confidence), and the other 15 events have strong preference for model S0 over model

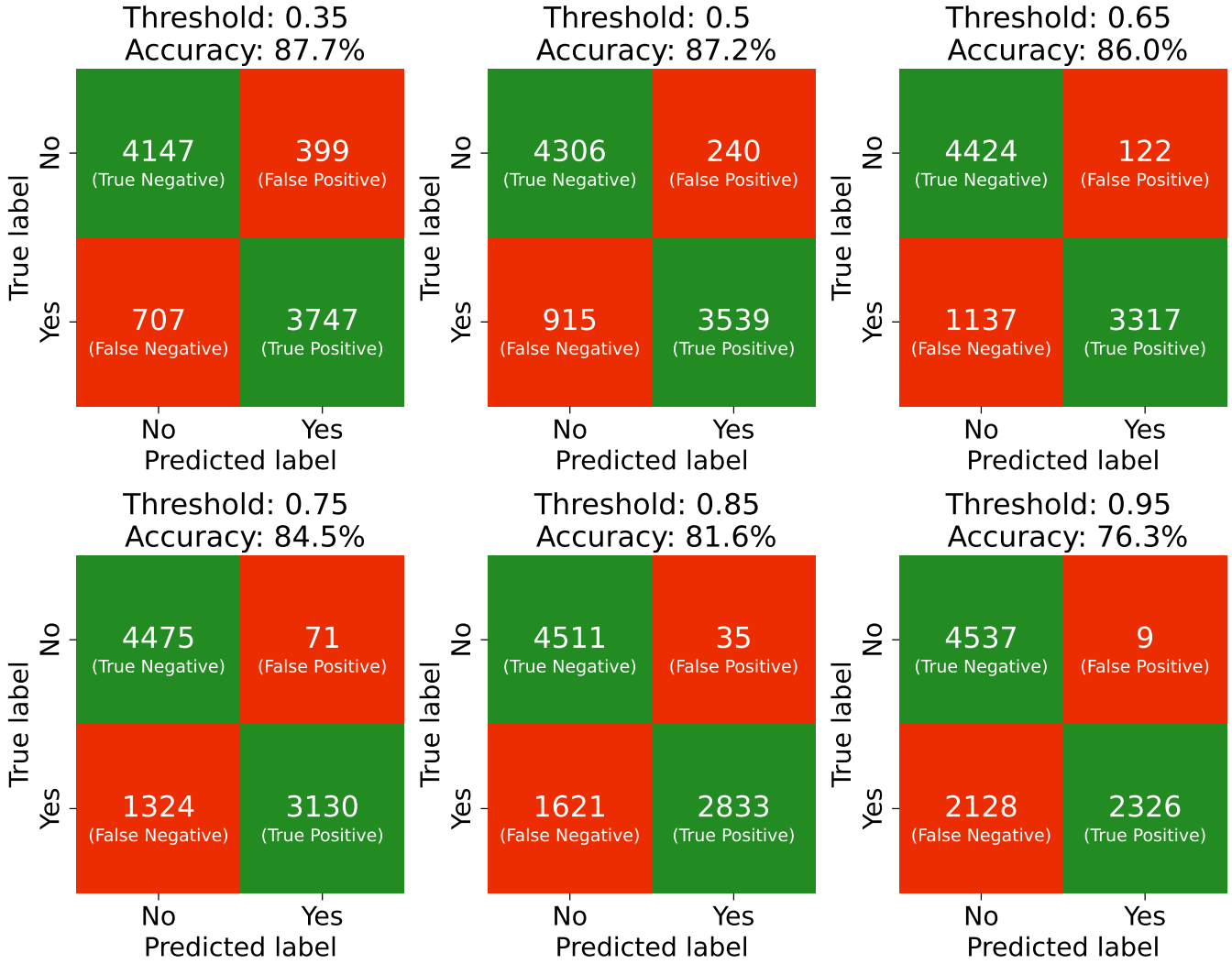

Figure 6. Confusion matrices build using the synthetic test dataset for 0.35, 0.5, 0.65, 0.75, 0.85, and 0.95 QPP-detection threshold values.

S1 and model S2 over model S1 (i.e., no QPP). The lightcurves of the selected flare events were downloaded from the GOES data archive, using GOES SSWIDL software[3].

For the considered sub-set of AFINO data, our FCN was unable to find any QPPs except one event. This is an expected result because of the above differences between the QPP properties in the AFINO catalogue and in our synthetic dataset. AFINO QPPs are weakly decaying which follows from the imposed PSD model (model S1). This makes it a different QPP type not coinciding with the QPP type considered in our study (long-period and rapidly decaying). In addition, our analysis showed that the QPP amplitudes in the AFINO catalogue are much smaller than those in our synthetic dataset, and AFINO QPPs appear across all flare phases, not only in the decay phase as we assumed in this work.

Thus, the developed FCN cannot be used for AFINO-type QPPs for the reasons mentioned above. However, this may be fixed in future by expanding the synthetic dataset used for the FCN training, to account for the QPP properties from the AFINO catalogue.

4.2. *Kepler data*

[3] https://hesperia.gsfc.nasa.gov/goes/goes.html

Table 2. The fraction of QPP found with the FCN in the Kepler stellar flare catalogue (Balona 2015) for different threshold values. A total of 2274 stellar flares were considered.

Threshold	0.35	0.5	0.65	0.75	0.85	0.95	0.99
QPP fraction, %	29.2	23.4	18.4	15.0	11.8	7.0	3.8

In contrast to the AFINO data, QPPs found in white-light stellar flares look similar to our dataset, even visually. To examine the performance of our FCN on this type of data, we used the set of 11 QPP events found by Pugh et al. (2016) in Kepler data with 1-min cadence by using the wavelet power spectrum of autocorrelation function, which show the stable decaying pattern. Figure 7 shows 7 out of 11 flare lightcurves from (Pugh et al. 2016) where the FCN found QPP, together with the QPP detection probability, $\mathcal{P}_{\rm QPP}$. The lightcurves where the FCN did not find QPP are presented in Fig. 8. This discrepancy with the results of (Pugh et al. 2016) seems natural. Indeed, comparing the lightcurves in Figs. 7 and 8, one may notice that the oscillatory patterns are more obvious (even visually) in Fig. 7. In contrast, the lightcurves in Fig. 8 where the FCN was not able to detect QPP in comparison with Pugh et al. (2016) seem to be more noisy and to have smaller relative amplitudes of possible QPP.

The FCN performance on QPPs found by Pugh et al. (2016) demonstrates that the FCN is able to detect QPPs in real lightcurves and can be used as a data sieve to detect, at least, the most obvious QPPs events. To proceed with this conclusion, we took the data from the stellar flare catalogue (Balona 2015) describing 3140 stellar flares observed with Kepler. The short cadence flare light curves were downloaded by their KIC and flare peak time from the flare catalogue using Lightkurve python library (Lightkurve Collaboration et al. 2018). Before using this data as the FCN input, we detrended the lightcurves using a parabolic fit, to get rid of the star rotation effect (see Fig. 9). In this fitting procedure, the second third of the time series where the flares are usually present was masked out so that the fit mainly captures the slowly-varying background trend. Next, we examined how well this slowly-varying trend best-fits the data using the χ^2-criterion, calculated at the first and last thirds of the time series. The left panel of Fig. 9 shows the histogram of the obtained χ^2 values. Using it, we empirically set our χ^2 threshold to 20, which accounts for most of the flare events and filters out the outliers. After filtering the lightcurves with $\chi^2 > 20$ and removing their slowly-varying background trends (see the middle and right panels of Fig. 9), we selected only flares with peaks exceeding the mean background flux value by 3 standard deviations, that gave us a final set of 2274 flare lightcurves to process with our FCN. The right panel of Fig. 2 demonstrates all these selected flare lightcurves plotted together after their min-max normalisation. It may be noted that we had to remove this slowly-varying background trend because it was not included in our synthetic lightcurves to train the network. However, in future studies, it can be added to the training dataset to avoid the need for detrending.

The obtained Kepler lightcurves were used as the FCN input to find QPPs. Table 2 presents the fraction of the lightcurves in which the FCN detected QPPs, for different QPP detection threshold values. For example, for $\mathcal{P}_{\rm QPP} > 0.5$ the fraction of Kepler flares with QPPs is found to be about 23%. However, for a more conservative $\mathcal{P}_{\rm QPP} > 0.95$ threshold, the QPP detection rate is about 7% which is comparable to the recent result of QPP detection in TESS flares (Ramsay et al. 2021) and is approximately factor of two higher than the previous estimation of the QPP detection rate in Kepler flares (Pugh et al. 2016; Balona 2015). Furthermore, these 7% detection rate of QPP in stellar white light flares seem to be very similar to the statistics of QPP in weak solar C-class flares (Hayes et al. 2020). Thus, we consider the performance of our FCN in this case very reasonable. For all the lightcurves where the FCN found QPP with $\mathcal{P}_{\rm QPP} > 0.95$ threshold (159 events in total), the star's KIC identifier and the BJD time of flare peak are summaries in Table 3 in Appendix A. This data may be used for a more detailed analysis of QPP, including with other methods, in future.

5. FCN INSTALLATION GUIDE AND USAGE EXAMPLE

The synthetic dataset used in this study and the source files for the developed FCN are available in open access via a GitHub repository[4] and a Harvard Dataverse dataset. The GitHub repository includes two Jupyter notebooks in Notebooks folder (one for synthetic dataset generation and one for the FCN training) and an easy to install and user-friendly Streamlit browser application, created by us for running the developed FCN. To use this application,

[4] github.com/BelovSA/QPP-Detection

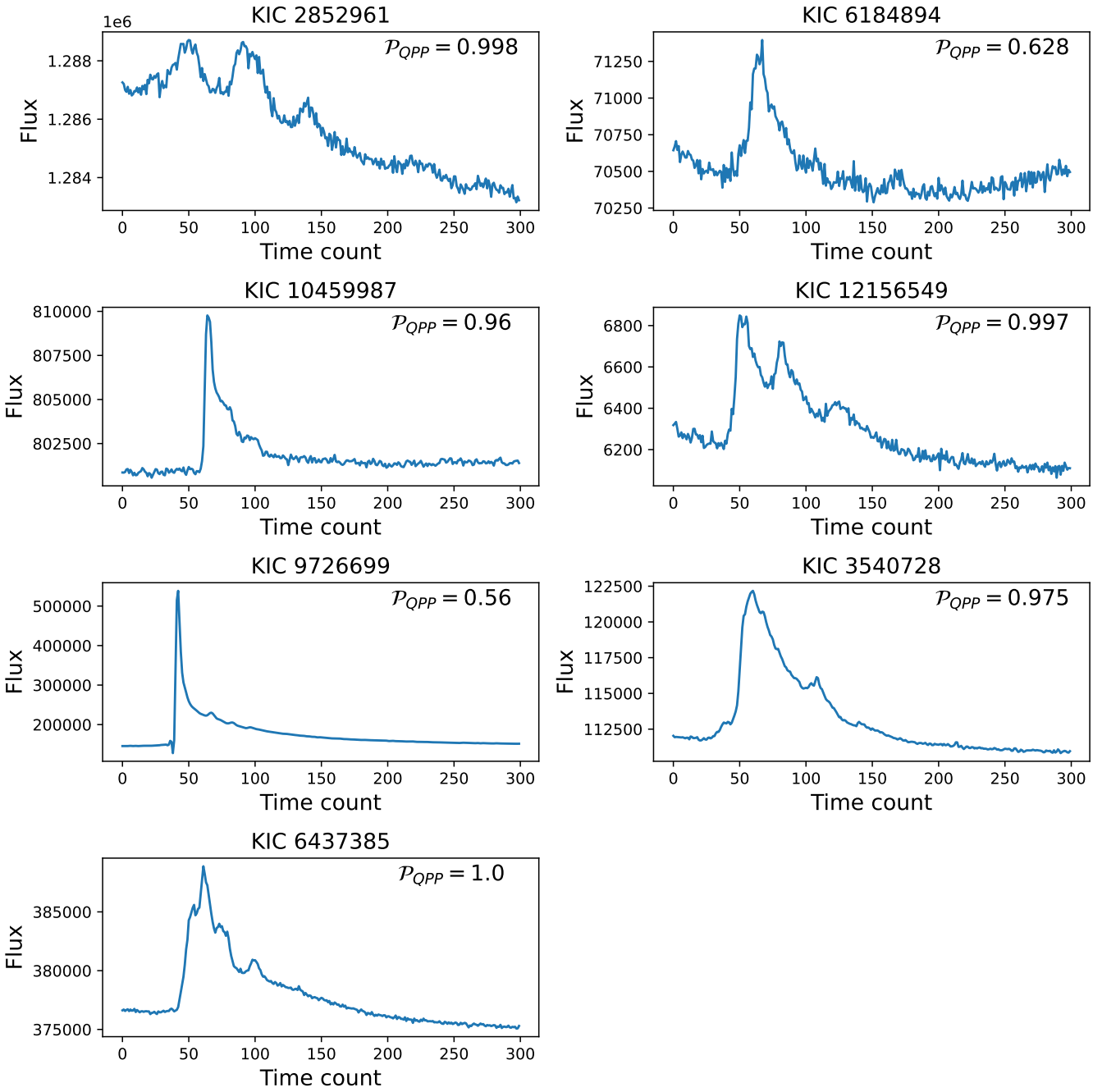

Figure 7. Stellar flare lightcurves with QPP from (Pugh et al. 2016) where the FCN also found QPP.

one should implement the following commands (written in Unix format, assuming Anaconda/Miniconda distribution v.24.1.2 or later has been pre-installed):

1. Clone or copy the project repository from GitHub.

2. In terminal, change your working directory to the root directory of the project:
   `cd QPP-Detection` (or `cd QPP-Detection-main`).

3. Create new Anaconda environment (can take several minutes):
   `conda env create -f ./Environment/env.yml`

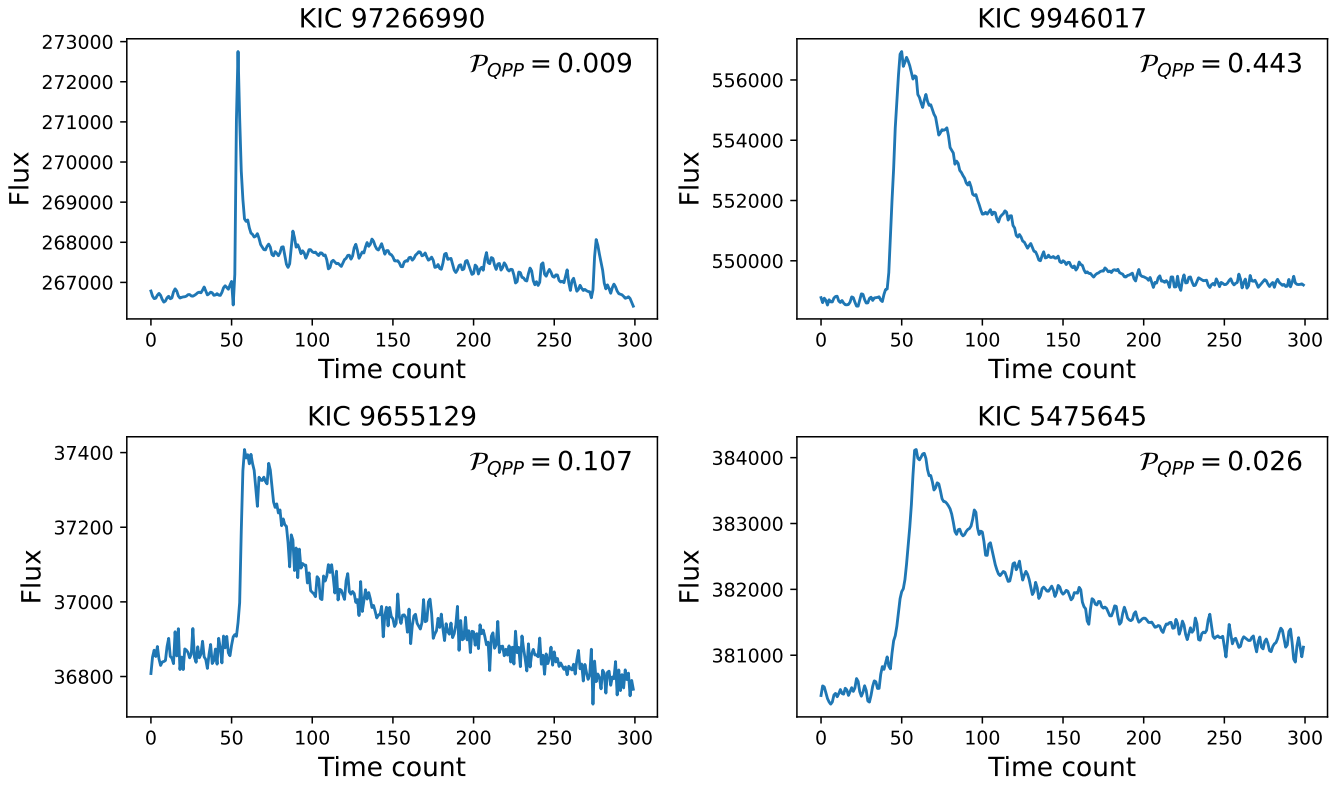

Figure 8. Stellar flare lightcurves with QPP from (Pugh et al. 2016) where the FCN did not find and QPP.

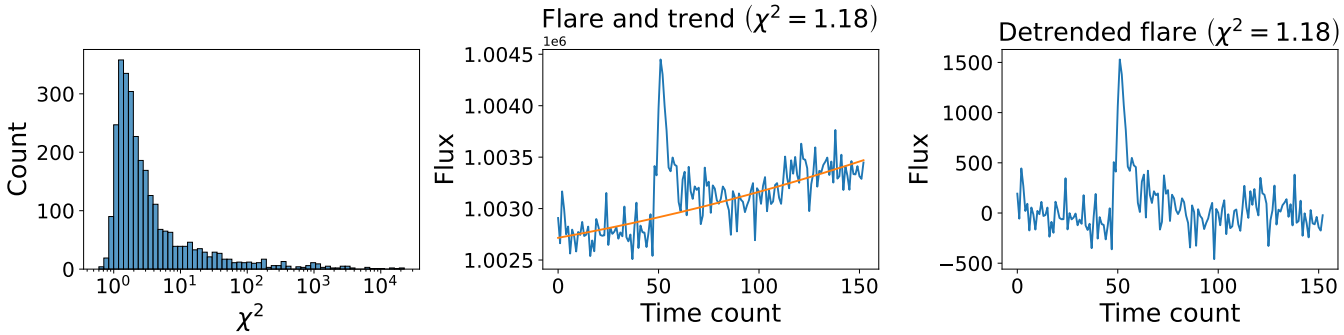

Figure 9. Left panel: the histogram of the χ^2-criterion characterising how good the slowly-varying background flare trend is fitted by a parabolic function, for stellar flare catalogue of Balona (2015) based on the white-light Kepler data. Middle panel: a flare light curve and its parabolic trend. Right panel: a flare light curve with the parabolic trend removed.

4. Activate the environment created:
   `conda activate qpp_detection`

5. Run the application[5]:
   `streamlit run ./Application/app.py`

[5] In case of trouble, the update of `streamlit` and `PyTorch-lightning` libraries may require:
`conda update <library_name>`

6. If the steps above are successful you can open the application (if it is not opened automatically) in your browser by the url:
http://localhost:8501 (or a similar one specified in terminal after running the previous command)

7. Use the graphical interface to detect QPP. The input should be `.csv` files with the lightcurve data in the column named 'flux' (other columns, if any, are ignored). The example lightcurves taken from (Pugh et al. 2016) and shown in Fig. 7 and Fig. 8 can be found in the "Test data" folder in the project directory.

Steps 1–3 should be done once to install the application, while steps 1 and 4–7 should be executed each time. The next subsection illustrates the use-case for the application. For the Windows computers, steps 1–7 should be done via Anaconda Prompt, not Windows shell.

Once steps 4–6 are done, the user should see the following window in a browser as shown in Fig. 10. This window initially consists of the search bar, where the folder with the data may be dragged and dropped, and the detect button. After clicking that button, the FCN processes the data and produces the output table containing the QPP probabilities and visualisations of the input flare profiles. Finally, the results may be downloaded by clicking the "Download data as CSV" button for further analysis.

6. CONCLUSIONS AND PROSPECTS

In this study we applied the FCN architecture proposed by Wang et al. (2017) to detect QPPs in light curves of solar and stellar flares. For this purpose, we generated 90,000 synthetic lightcurves using the flare models by Davenport et al. (2014), Gryciuk et al. (2017), and a guessed flare model made by two half-Gaussians (Broomhall et al. 2019). Then, decaying QPPs were added to the decay phase of 50% of the lightcurves generated. Also, noise (white, red or both) was added to all the lightcurves. The obtained dataset was divided into train, validation and test datasets.

To train the FCN, we used two input channels: an original lightcurve and a detrended lightcurve with a rise phase padded by white noise. After training, the FCN showed an accuracy of 87.2% on the test data with a precision score of 93.6%. The change of the detection threshold from 0.5 to 0.95 led to a decrease in a number of false positive detections, as well as, to a decrease in the accuracy score making the FCN more conservative.

After testing the FCN performance on the synthetic data, we applied the FCN to real data examples. For the 30 selected samples from the AFINO catalogue (Inglis et al. 2015, 2016) 50% of which contain QPP and 50% do not, the FCN was able to find only one QPP. The reason for this is that the type of QPP in the AFINO catalogue is apparently different from that in our training dataset. To test the FCN on data similar to the training data, we used 11 lightcurves of stellar flares with strong evidence of decaying QPP presence found by Pugh et al. (2016) in Kepler observations. For this set, the FCN detected QPPs in 7 out of 11 flares. The remaining 4 lightcurves had either noisy profiles or small QPP amplitudes.

To proceed with Kepler data, we used the information about 3140 lightcurves from the stellar flare catalogue created by Balona (2015). We detrended these lightcurves using a slowly-varying parabolic fit and filtered out the lightcurves where the fit was unsuccessful using the χ^2-criterion. Additionally, we selected the samples where the peak amplitude exceeded the mean by 3σ. After this preparation, 2274 lightcurves were fed to the FCN. For the threshold $\mathcal{P}_{\rm QPP} > 0.95$, the FCN detected QPPs in 7% of all flare lightcurves considered. This is comparable with the QPP detection rate in stellar flares, estimated recently by Ramsay et al. (2021).

We made our synthetic dataset and the source code available in open access via Harvard Dataverse and GitHub. Our GitHub repository has two Jupyter notebooks to generate the dataset and train the FCN with it. It also contains the Streamlit browser application, developed by us, which allows one to run the pre-trained FCN and use it for QPP detection in future studies. We also provided a detailed user guide for the installation, running, and using the developed FCN.

We used the FCN time series classifier, based on the architecture proposed in (Wang et al. 2017), as the first step in applying Deep Learning techniques for the QPP detection task. As for the next step, a more complicated network architecture can be used including different approaches such as residual (ResNet) and recurrent (RNN) networks. However, this approach demands powerful hardware and may limit the inference speed that can be crucial for large-scale surveys. Another possible development is data engineering. The additional features may be used as the new FCN input channels. For example, the autocorrelation function can be calculated for the detrended lightcurve and passed to the FCN as an additional channel. Finally, the raw lightcurves can be transformed into 2D spectra (via, for example, the wavelet transform) and then examined by the image recognition techniques for the presence of characteristic QPP signatures.

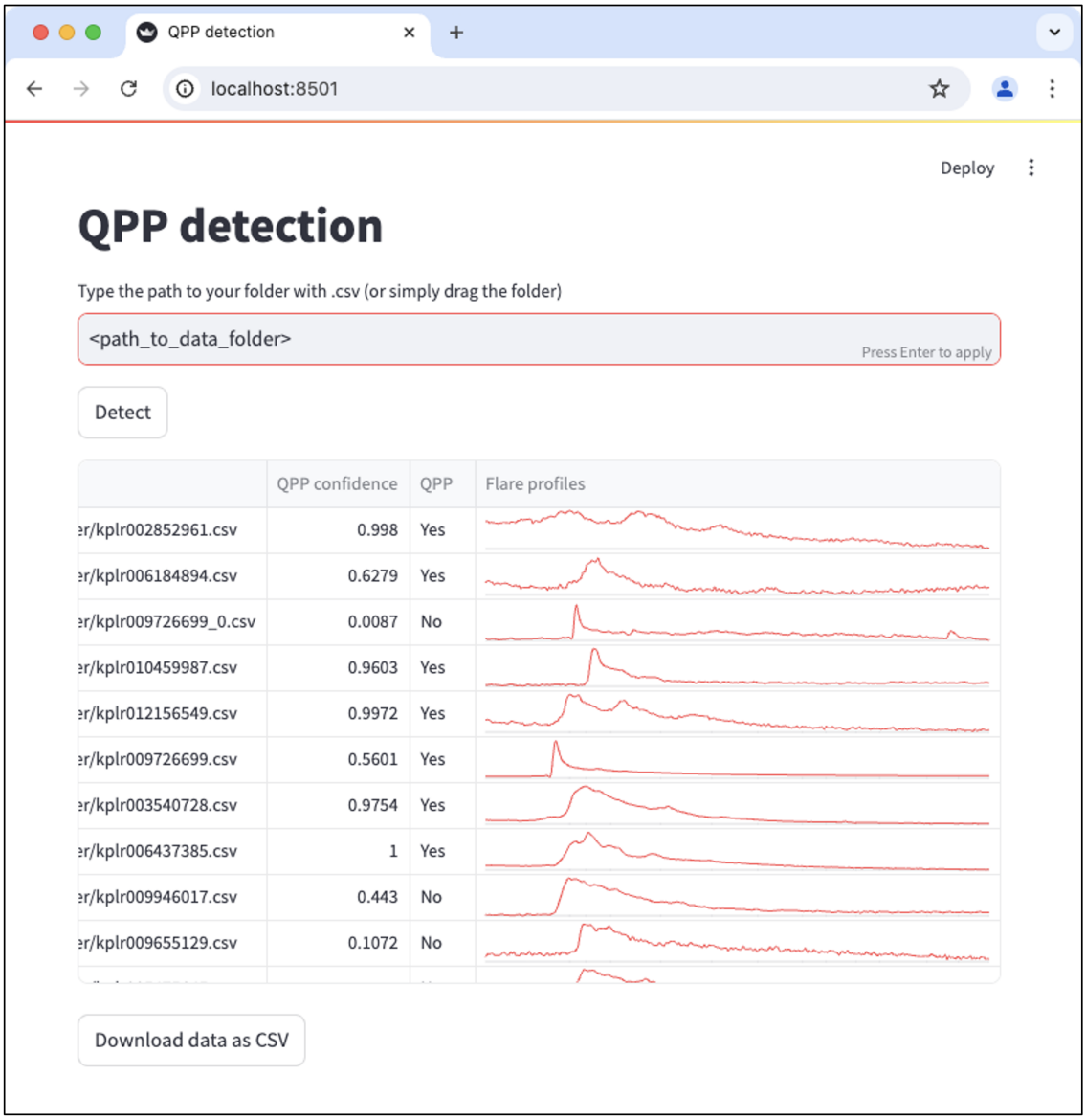

Figure 10. The graphical interface of the browser application for the FCN-based QPP detection in the flare lightcurves data.

The work is funded by the STFC Grant ST/X000915/1. DYK and VMN also acknowledge funding from the Latvian Council of Science Project No. lzp2022/1-0017.

Software: NumPy, a Python packge for fundamental scientific computing (Harris et al. 2020); SciPy, a Python package for fundamental algorithms for scientific computing (Virtanen et al. 2020); Lightkurve, a Python package for Kepler and TESS data analysis (Lightkurve Collaboration et al. 2018); Astopy, a Python package for astrophysical purposes (Astropy Collaboration et al. 2013, 2018, 2022); PyTorch Lightning, the deep learning framework; Streamlit, a Python library for deploying ML projects.

APPENDIX

Table 3. KIC and BJD flare peak times for 159 flares with $\mathcal{P}_{\rm QPP} \geq 0.95$ found in the flare catalogue (Balona 2015).

KIC	BJD	KIC	BJD	KIC	BJD	KIC	BJD
11560431	2456115.361	10160534	2455064.852	11560431	2456173.104	10355856	2455387.912
3430868	2455038.978	5357275	2455119.591	11551430	2455013.017	11551430	2455029.862
3239945	2456170.505	5733906	2455099.532	4671547	2455073.743	11560431	2456056.445
11560431	2456072.751	11610797	2454981.631	11551430	2456131.799	11665620	2455806.894
11548140	2455386.295	11551692	2456170.504	5475645	2456330.43	6106152	2455024.179
5357275	2455106.976	3441906	2455952.2	9349698	2456354.714	11709006	2455408.873
9895004	2455437.151	11551692	2456042.642	10459987	2456179.644	8429280	2455004.753
5557932	2455038.978	11189959	2455114.891	11560431	2456110.758	11560431	2456197.291
6106152	2455018.608	9652680	2455090.13	1161345	2456170.506	10063343	2455151.39
11560431	2456172.309	11231334	2455491.728	5108214	2455837.489	4273689	2455249.467
11560431	2455020.996	11709006	2454971.877	4758595	2456233.845	11551430	2455024.522
6548447	2455323.093	7940546	2455768.88	11551430	2456194.753	3239945	2456156.116
4543412	2455163.976	11560431	2456117.885	4831454	2455187.025	9761199	2455799.706
8651471	2455845.632	7339343	2455005.364	7206837	2456084.741	11551430	2456279.936
11709006	2455430.144	9641031	2456219.273	11560431	2456064.381	8429280	2455032.148
8429280	2455006.133	11560447	2455055.708	11709006	2454971.707	5609753	2456399.539
11560431	2455013.574	10528093	2456262.771	4758595	2456226.538	7841024	2454968.701
11548140	2455486.206	9821078	2455623.573	11560431	2455023.444	12644769	2456170.501
11551692	2456289.977	9641031	2455022.582	11548140	2455488.308	11665620	2455791.997
2300039	2455773.901	2162635	2455806.895	12156549	2455346.612	11560431	2456057.61
12156549	2455342.383	9641031	2456157.95	7940546	2455779.026	11560431	2456165.778
4758595	2456210.581	12156549	2455376.074	11551430	2456245.555	2302548	2456170.505
7940546	2455871.277	4568729	2455056.284	9705459	2456371.072	9641031	2456109.318
1025986	2455133.355	11560431	2456117.449	9655129	2456148.63	9655129	2456112.973
4671547	2455087.968	11560431	2456150.678	11709006	2454969.451	11551430	2456217.422
11560431	2456184.967	5557932	2455039.036	4671547	2455074.027	12156549	2455347.197
5522786	2456011.854	11560431	2456070.227	12102573	2455072.339	11560431	2455003.17
4939265	2456281.794	4671547	2455090.038	6286925	2455038.42	9641031	2456100.371
2300039	2455790.849	3430868	2455039.036	7765135	2455072.194	11551430	2456170.457
1161345	2455806.895	5522786	2455871.3	11560431	2456194.578	12156549	2455318.629
11551430	2456296.8	1871056	2456156.115	6106152	2455013.524	7692454	2456266.216
8429280	2455032.246	10355856	2454982.87	9821078	2455705.652	7940546	2456094.427
9833666	2456394.052	6442183	2455860.26	11231334	2456178.25	6205460	2455334.886
11560431	2455010.158	7940546	2456259.122	12102573	2455079.331	11551430	2456235.537
7206837	2455806.894	11560431	2456186.81	11560431	2456170.503	11560431	2456114.382
12156549	2455086.645	9833666	2456423.294	11560431	2456156.548	7940546	2455048.776
11560431	2456112.473	11560431	2456142.847	7940546	2455503.213	6106152	2455033.037
3128488	2454964.786	12156549	2455067.352	4273689	2455263.655	11709006	2454966.828
11231334	2455435.206	12156549	2455374.593	11548140	2455453.673	11551430	2456278.089
12418816	2455257.868	11560431	2456132.526	4758595	2456219.141		

A. KEPLER FLARE WITH QPP

REFERENCES

Ahmed, O. W., Qahwaji, R., Colak, T., et al. 2013, SoPh, 283, 157, doi: 10.1007/s11207-011-9896-1

Anfinogentov, S. A., Antolin, P., Inglis, A. R., et al. 2022, SSRv, 218, 9, doi: 10.1007/s11214-021-00869-w

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33, doi: 10.1051/0004-6361/201322068

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123, doi: 10.3847/1538-3881/aabc4f

Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, ApJ, 935, 167, doi: 10.3847/1538-4357/ac7c74

Bai, J.-Y., Wang, J., Li, H. L., et al. 2023, PASP, 135, 064201, doi: 10.1088/1538-3873/acdea8

Balona, L. A. 2015, Monthly Notices of the Royal Astronomical Society, 447, 2714–2725, doi: 10.1093/mnras/stu2651

Benz, A. O. 2017, Living Reviews in Solar Physics, 14, 2, doi: 10.1007/s41116-016-0004-3

Bobra, M. G., & Couvidat, S. 2015, ApJ, 798, 135, doi: 10.1088/0004-637X/798/2/135

Broomhall, A.-M., Davenport, J. R. A., Hayes, L. A., et al. 2019, The Astrophysical Journal Supplement Series, 244, 44, doi: 10.3847/1538-4365/ab40b3

Camporeale, E. 2019, Space Weather, 17, 1166, doi: 10.1029/2018SW002061

Cho, I. H., Cho, K. S., Nakariakov, V. M., Kim, S., & Kumar, P. 2016, ApJ, 830, 110, doi: 10.3847/0004-637X/830/2/110

Davenport, J. R. A., Hawley, S. L., Hebb, L., et al. 2014, The Astrophysical Journal, 797, 122, doi: 10.1088/0004-637x/797/2/122

Fluke, C. J., & Jacobs, C. 2020, WIREs Data Mining and Knowledge Discovery, 10, e1349, doi: 10.1002/widm.1349

Georgoulis, M. K., Bloomfield, D. S., Piana, M., et al. 2021, Journal of Space Weather and Space Climate, 11, 39, doi: 10.1051/swsc/2021023

Gryciuk, M., Siarkowski, M., Sylwester, J., et al. 2017, Solar Physics, 292, doi: 10.1007/s11207-017-1101-8

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357, doi: 10.1038/s41586-020-2649-2

Hayes, L. A., Inglis, A. R., Christe, S., Dennis, B., & Gallagher, P. T. 2020, ApJ, 895, 50, doi: 10.3847/1538-4357/ab8d40

Inglis, A. R., Ireland, J., Dennis, B. R., Hayes, L., & Gallagher, P. 2016, The Astrophysical Journal, 833, 284, doi: 10.3847/1538-4357/833/2/284

Inglis, A. R., Ireland, J., & Dominique, M. 2015, The Astrophysical Journal, 798, 108, doi: 10.1088/0004-637x/798/2/108

Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L., & Muller, P.-A. 2019, Data Mining and Knowledge Discovery, 33, 917–963, doi: 10.1007/s10618-019-00619-1

Jeong, H.-J., Moon, Y.-J., Park, E., & Lee, H. 2020, ApJL, 903, L25, doi: 10.3847/2041-8213/abc255

Kim, S., Nakariakov, V. M., & Shibasaki, K. 2012, ApJL, 756, L36, doi: 10.1088/2041-8205/756/2/L36

Kingma, D. P., & Ba, J. 2014, Adam: A Method for Stochastic Optimization, arXiv, doi: 10.48550/ARXIV.1412.6980

Kolotkov, D. Y., Pugh, C. E., Broomhall, A.-M., & Nakariakov, V. M. 2018, ApJL, 858, L3, doi: 10.3847/2041-8213/aabde9

Kupriyanova, E., Kolotkov, D., Nakariakov, V., & Kaufman, A. 2020, Solar-Terrestrial Physics, 6, 3, doi: 10.12737/stp-61202001

Lightkurve Collaboration, Cardoso, J. V. d. M., Hedges, C., et al. 2018, Lightkurve: Kepler and TESS time series analysis in Python, Astrophysics Source Code Library. http://ascl.net/1812.013

McLaughlin, J. A., Nakariakov, V. M., Dominique, M., Jelínek, P., & Takasao, S. 2018, SSRv, 214, 45, doi: 10.1007/s11214-018-0478-5

Mehta, T., Broomhall, A. M., & Hayes, L. A. 2023, MNRAS, 523, 3689, doi: 10.1093/mnras/stad1619

Nakariakov, V. M., Anfinogentov, S., Storozhenko, A. A., et al. 2018, ApJ, 859, 154, doi: 10.3847/1538-4357/aabfb9

Nakariakov, V. M., Inglis, A. R., Zimovets, I. V., et al. 2010, Plasma Physics and Controlled Fusion, 52, 124009, doi: 10.1088/0741-3335/52/12/124009

Nakariakov, V. M., Kolotkov, D. Y., Kupriyanova, E. G., et al. 2019a, Plasma Physics and Controlled Fusion, 61, 014024, doi: 10.1088/1361-6587/aad97c

Nakariakov, V. M., Kosak, M. K., Kolotkov, D. Y., et al. 2019b, ApJL, 874, L1, doi: 10.3847/2041-8213/ab0c9f

Nakariakov, V. M., & Melnikov, V. F. 2009, SSRv, 149, 119, doi: 10.1007/s11214-009-9536-3

Nishizuka, N., Sugiura, K., Kubo, Y., Den, M., & Ishii, M. 2018, ApJ, 858, 113, doi: 10.3847/1538-4357/aab9a7

Pascoe, D. J., Smyrli, A., Van Doorsselaere, T., & Broomhall, A. M. 2020, ApJ, 905, 70, doi: 10.3847/1538-4357/abc69d

Pugh, C. E., Armstrong, D. J., Nakariakov, V. M., & Broomhall, A.-M. 2016, Monthly Notices of the Royal Astronomical Society, 459, 3659–3676, doi: 10.1093/mnras/stw850

Ramsay, G., Kolotkov, D., Doyle, J. G., & Doyle, L. 2021, SoPh, 296, 162, doi: 10.1007/s11207-021-01899-x

Reale, F., Testa, P., Petralia, A., & Kolotkov, D. Y. 2019, ApJ, 884, 131, doi: 10.3847/1538-4357/ab4270

Shibata, K., & Magara, T. 2011, Living Reviews in Solar Physics, 8, 6, doi: 10.12942/lrsp-2011-6

Simões, P. J. A., Hudson, H. S., & Fletcher, L. 2015, SoPh, 290, 3625, doi: 10.1007/s11207-015-0691-2

Tilipman, D., Kazachenko, M., Tremblay, B., et al. 2023, ApJ, 956, 83, doi: 10.3847/1538-4357/ace621

Valach, F., Revallo, M., Bochníček, J., & Hejda, P. 2009, Space Weather, 7, S04004, doi: 10.1029/2008SW000421

Vida, K., & Roettenbacher, R. M. 2018, A&A, 616, A163, doi: 10.1051/0004-6361/201833194

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261, doi: 10.1038/s41592-019-0686-2

Wang, T. 2011, SSRv, 158, 397, doi: 10.1007/s11214-010-9716-1

Wang, T., Ofman, L., Yuan, D., et al. 2021, SSRv, 217, 34, doi: 10.1007/s11214-021-00811-0

Wang, Y., Liang, B., & Feng, S. 2024, ApJS, 270, 4, doi: 10.3847/1538-4365/ad09b6

Wang, Y., Liu, J., Jiang, Y., & Erdélyi, R. 2019, ApJ, 881, 15, doi: 10.3847/1538-4357/ab2b3e

Wang, Z., Yan, W., & Oates, T. 2017, in 2017 International Joint Conference on Neural Networks (IJCNN) (IEEE), doi: 10.1109/ijcnn.2017.7966039

Webbe, R., & Young, A. J. 2023, RAS Techniques and Instruments, 2, 238, doi: 10.1093/rasti/rzad015

Xu, C., Wang, J. T. L., Wang, H., et al. 2024, SoPh, 299, 36, doi: 10.1007/s11207-024-02283-1

Zhang, T., Hao, Q., & Chen, P. F. 2024, ApJS, 272, 5, doi: 10.3847/1538-4365/ad3039

Zimovets, I. V., McLaughlin, J. A., Srivastava, A. K., et al. 2021, SSRv, 217, 66, doi: 10.1007/s11214-021-00840-9