Feature Identification and Statistical Characteristics of Quasi-periodic Pulsation in Solar Flares using the Markov-Chain-Monte-Carlo Approach

Yangfan Guo[1], Bo Liang[1], Song Feng[1], Ding Yuan[2], Valery M. Nakariakov[3,4], Wei Dai[1], and Yunfei Yang[1]
[1] Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, People's Republic of China
feng.song@kust.edu.cn
[2] Institute of Space Science and Applied Technology, Harbin Institute of Technology, Shenzhen, Guangdong 518055, People's Republic of China
yuanding@hit.edu.cn
[3] Centre for Fusion, Space and Astrophysics, Department of Physics, University of Warwick, CV4 7AL, Coventry, UK
[4] Centro de Investigacion en Astronomía, Universidad Bernardo O'Higgins, Avenida Viel 1497, Santiago, Chile

Abstract

Quasi-periodic pulsation (QPP) is a common phenomenon in solar flares. Studying QPP is important to further our understanding of the physical processes operating in flares. However, detection of QPP is complicated by the presence of noise in flaring lightcurves. In this study, we apply the Bayesian-based Markov-Chain-Monte-Carlo (MCMC) technique to the QPP detection. We use MCMC to fit the Fourier power spectral density (PSD) profiles of flaring lightcurves, aiming to determine a quasi-periodic component by model comparison and test statistics. Two models fitting the PSD were compared: the first model consists of colored and white noise only, and the second model adds a spectral peak of a Gaussian shape representing a short-living oscillatory signal. To evaluate MCMC of the QPP detection, we test it on 100 synthetic signals with spectral properties similar to those observed in flares. Subsequently, we analyzed QPP events in 699 flare signals in the 1–8 Å channel recorded by the Geostationary Operational Environmental Satellite from 2010 to 2017, including 250 B-class, 250 C-class, 150 M-class, and 49 X-class flares. Approximately 57% X-class, 39% M-class, 20% C-class, and 16% B-class flares are found to show a strong evidence of QPP, whose periods range mainly from 6.2 to 75.3 s. The results demonstrate that QPP events are easier to detect in more powerful flares. The distribution of the detected QPP periods is found to follow a logarithmic normal distribution. The distributions in the four flare classes are similar. This suggests that the established distribution is a common feature for flares of different classes.

Unified Astronomy Thesaurus concepts: Solar flares (1496); Solar flare spectra (1982); Solar activity (1475); Solar oscillations (1515)

1. Introduction

In the past few decades, many studies have proved that quasi-periodic pulsation (QPP) is a common feature of solar flares (e.g., Nakariakov et al. 2010; Inglis et al. 2016; Tian et al. 2016; Hayes et al. 2020; Zimovets et al. 2021). Typically, QPP is defined as a sequence of pluses or cyclic variations of the emission, with similar (while not necessarily identical) time intervals between the successive peaks or time durations of the oscillation cycles. QPP events have been observed in all phases of a flare (i.e., in the pre-flare, impulsive, and decay phases), and in both thermal and non-thermal emissions (e.g., Zimovets & Struminsky 2009; Tan & Tan 2012; Milligan et al. 2017; Clarke et al. 2021).

Typical QPP periods range from a few seconds to several minutes (Nakariakov 2007; Nakariakov & Melnikov 2009; Yuan et al. 2019; Li et al. 2020a, 2020b). The vast majority of QPP studies have focused on flares of powerful M- and X-classes, or some high-power C-class flares (Inglis et al. 2016; Kolotkov et al. 2018; Hayes et al. 2019; Yuan et al. 2019). QPP has also been detected in flares of weaker classes (Nakariakov et al. 2018; Hayes et al. 2020). Often, QPP appears to have a non-stationary character, i.e., the instantaneous period, amplitude, and signal shape seem to vary in time (e.g., Nakariakov et al. 2019).

However, whether the apparent duration of a QPP pattern reflects its true duration or is limited by observational conditions is unclear. Furthermore, whether a QPP occurs in a certain time interval in a certain phase of a flare or if the actual duration of the QPP process is comparable with the whole duration of the flare or is even longer is also unclear. The occasional detection of QPP in the pre-flare phase (e.g., Tan et al. 2016; Abramov-Maximov & Bakunina 2020; Li et al. 2020b; Altyntsev et al. 2022; and Zimovets et al. 2022 for an alternative point of view) could be considered in favor of the latter scenario, at least for certain classes of QPP.

Most QPP studies have addressed one or a few specific events. However, statistical studies of QPP that have addressed the distribution of their parameters, the relationships between them and the parameters of the host flare, and the detection confidence levels have also been conducted. The most comprehensive investigations have been carried out by Inglis et al. (2016), who analyzed 675 M- and X-class flares observed by the Geostationary Operational Environmental Satellite (GOES) and 261 events from Fermi/Gamma Burst Monitor, and Hayes et al. (2020), who analyzed 5516 X-, M-, and C-class flares observed with GOES from 2011 to 2018. In addition, a number of smaller-scale statistical studies have addressed specific types of QPP or flares (e.g., Kupriyanova et al. 2010; Simões et al. 2015; Pugh et al. 2017b).

Together with the non-stationary character, QPP detection is also complicated by the presence of colored and white noise in the data (e.g., Vaughan 2005, 2010; Inglis et al. 2015, 2016;

Pugh et al. 2017a; Liang et al. 2020). Colored noise is characterized by a power abundance at low-frequency spectral components, while white noise is characterized by a power abundance that is uniformly distributed over all frequencies. Thus, it is important to estimate the effect of colored and white noise on the detection and analysis of QPP signals. In particular, some standardization of the QPP detection procedure (i.e., the minimization of the subjective choices of various parameters) is highly demanded.

The most straightforward method to detect QPP, based on the assumption of the harmonic and long-living nature of a QPP signal, is to estimate the significance of a candidate spectral peak above the noise level in the time frequency domain using signal processing techniques such as Fourier transform, periodograms, or wavelet transform (Broomhall et al. 2019).[5] The Bayesian-based Markov-Chain-Monte-Carlo (MCMC) technique is often used as the fitting method to estimate the most probable period of a QPP signal (Inglis et al. 2015, 2016; Sharma 2017; Hayes et al. 2020; Anfinogentov et al. 2021).

In the present work, we investigate distributions of oscillation periods of QPP detected in the thermal emission generated by solar flares, and then search for relationship between the period distributions and the host flare classes. This study is based on an analysis of the power spectra of flaring lightcurves. The feasibility of the robustness of the QPP detection and the oscillation period estimation with Bayesian-based MCMC is verified by its application to 100 synthetic QPP signals. Each synthetic signal included a quasi-periodic component, colored noise, and white noise (see Broomhall et al. 2019). In general, a flaring lightcurve also has a trend that can either be removed by smoothing or averaging, or be present in the signal. In the latter case, the power spectrum would resemble the colored noise spectrum. In the processing, two different models—with and without a QPP—are used to fit the power spectral density (PSD) as a function of frequency. The technique was then applied to randomly selected 699 solar flare signals that were recorded by GOES from 2010 to 2017, including X-, M-, C-, and B-class flares. We estimated the mean oscillation periods using Bayesian-based MCMC and then explored the statistical characteristics of QPP in the flares.

The rest of this paper is structured as follows. In Section 2, we describe the methodology for the QPP detection, including Bayesian-based MCMC, model comparison and calculating statistics. Sections 3 and 4 present the results of the analysis of the synthetic signals and solar flare signals. A discussion of the results and a conclusion are given in Section 5.

2. Methodology

2.1. Data Preprocessing

The first step of the analysis is to normalize an input signal,

$$S_{\mathrm{norm}} = \frac{S - \overline{S}}{\overline{S}}, \qquad (1)$$

where S is the original signal and $\overline{S}$ is the mean of the signal. We then calculate the PSD of the normalized input signal S_{norm} multiplied by a time window function, to minimize the effects of the finite duration of the signal. In our case, because the

input signal is evenly sampled, the PSD is constructed with the use of the standard discrete Fourier transform (i.e., a periodogram). In this step, we tested several window functions (e.g., the Blackman-, triangular-, and Hanning-window functions) to evaluate the effect of the frequency leakage. We found that the accuracy of the frequency representation using the Hanning window is superior to the other window functions, giving good frequency resolution and less frequency leakage. In the following calculations, we apply the Hanning-window function.

2.2. Bayesian-based MCMC

Bayesian-based MCMC is a sampling method that performs well with multi-parameter distributions (Hastings 1970; Anfinogentov et al. 2021). In the context of QPP detection, it allows us to fit a PSD with a certain guessed function (i.e., a model), determine its best-fitting parameters, and then compare different models. The Bayes theorem states

$$p(\Theta|D, M) = \frac{p(\Theta|M)p(D|\Theta, M)}{p(D|M)}, \qquad (2)$$

where D is the PSD of an observed signal, M denotes a selected model and $\Theta (= \{\Theta_1, \Theta_2, \ldots\})$ presents the parameters of model M. The quantity $p(D|\Theta, M)$ is the likelihood, which represents the probability of obtaining the observed data under the assumption of the model. The quantity $p(\Theta|M)$ is the prior distribution of the parameters, which indicates the estimation of parameters and is based on existing information. The coefficient $p(D|M)$ is a normalizing constant. The quantity $p(\Theta|D, M)$ is the posterior distribution of the parameters, which is the re-estimation of the probability of the parameter based on the observed data. Therefore, the posterior distribution is represented by the product of the prior distribution and the likelihood.

The periodogram of any observed stochastic time series with the regular time axis of length N has the Fourier power $D_j^{\mathrm{obs}}(= \{D_1^{\mathrm{obs}}, \ldots, D_{\frac{N}{2}}^{\mathrm{obs}}\})$ at a Fourier frequency $f_j (= \{f_1, \ldots, f_{\frac{N}{2}}\})$, and is exponentially distributed about the true spectral density $S_j = S(f_j)$ (Chatfield 2003). So the likelihood can be written as

$$p(D|\Theta, M) = \prod_{j=1}^{\frac{N}{2}} \frac{1}{S_j} \exp\left(-\frac{D_j}{S_j}\right). \qquad (3)$$

In the Bayesian technique, the choice of the prior affects the posterior distribution of the parameters. Here, each parameter distribution is considered as a normal distribution because each parameter is affected by each independent sampling point. In a previous study, Vaughan (2010) set different priors and then observed the influence on the posterior, and came to the conclusion that the independent normal distribution results for parameter setting were the most appropriate. Thus, we set the parameters of the selected model as an independent normal distribution.

After determining the prior distribution and likelihood, we use MCMC to sample the posterior distribution. MCMC can effectively map a posterior distribution in a multi-dimensional parameter space. After some initial iterations (known as burn-in), the values of each parameter of the selected model will converge and the Markov chain returns samples of convergent parameter values. The convergent distribution of the Markov

---

[5] For alternative QPP detection techniques, see, e.g., Kolotkov et al. (2015), Broomhall et al. (2019), Anfinogentov et al. (2022).

chain is the same as the posterior distribution, which means that the parameters of the selected model are frequently sampled within a more reasonable range. MCMC provides us with the sampling results of each parameter of the selected model, which can be used to calculate the parameters of the best-fit model and for the subsequent statistical calculations.

2.3. Model Selection and Model Comparison

We selected two candidate models to fit PSD of the analyzed signals. The first model, M_0, is a power law plus a constant model,

$$M_0 = Af^{-\alpha} + c, \tag{4}$$

where $Af^{-\alpha}$ with constant A and $\alpha > 0$ is used to reflect the Fourier power-law property of colored noise, and constant c is to reflect the constant property of white noise.

The second model, M_1, is the power-law-plus-constant model with an additional Gaussian component around a certain frequency, which accounts for the possible appearance of the QPP in the solar flare, i.e.,

$$M_1 = M_0 + B \exp\left\{-\frac{(\log(f) - \beta)^2}{2\sigma^2}\right\}, \tag{5}$$

where B is the amplitude of the Gaussian component, β is its central frequency of the Gaussian component, and σ is the width of the Gaussian. The finite width of the spectral peak reflects the finite duration of the QPP signal. It could also account for the transient, non-stationary nature of QPP patterns (Nakariakov et al. 2019). To a given model, each parameter distribution is adjusted according to different input data, including their mean and standard deviation.

To determine which model is more consistent with the PSD of a given signal, we use the Bayesian information criterion (BIC, Schwarz 1978; Burnham & Anderson 2004; Anfinogentov et al. 2021), which is defined as

$$BIC = -2\log(L) + k\log(n), \tag{6}$$

where L is the maximum likelihood of the fit, k is the number of free parameters of the selected model, and n is the number of data points used for fitting. The smaller the result calculated by this method, the better the fitting result. BIC takes into account the number of parameters of the model. The model with more parameters will be punished to reduce the probability of an over-fitting model being selected. By comparing the BIC of two fitted models, we choose the model that is more suitable for the PSD of the input signal, calculating the difference of BICs of the models as

$$\Delta BIC = BIC_{M_0} - BIC_{M_1}, \tag{7}$$

where the indices correspond to the compared models. A positive ΔBIC would imply that model M_1 is superior to model M_0, otherwise model M_0 is better. In this step, we take $|\Delta BIC| \geqslant 10$ as the criterion that one model is significantly superior to another (Burnham & Anderson 2004).

It is worth noting that BIC is only used to measure which of the two compared models is more suitable for the input data. It does not confirm that any of them is a good choice. It is possible that both models are not very suitable. After all, for any data, we can use countless number of models to get

countless fitting results. To confirm how well a fitted model matches the data, we need to calculate the test statistics.

2.4. Test Statistic and Significance Testing

The statistic that we use is as follows (Vaughan 2010).

$$T_R = \max_j \left(\frac{2D_j}{M_j}\right), \tag{8}$$

where the index j ranges from 1 to $N/2$. The quantity T_R reflects the maximum deviation between the observed value D_j and the model M_j; that is, T_R indicates the local anomalies between two data sets.

As mentioned in Section 2.2, by sampling the Bayesian posterior distribution with the use of MCMC, we obtain sampling results of each parameter of the selected model M_0 or M_1. For either M_0 or M_1, after a certain number of iterations, the values of every parameter of the selected model will converge. If we set the sampling times after convergence to K, then we will get K sets of model parameters, which are used to generate K groups of simulated model data $B_i (= \{B_1, \ldots, B_K\})$. In each group B_i, we multiply each datum B_{ij} by a random number X_{ij} from a chi-square distribution with two degrees of freedom (Vaughan 2010), and obtain K groups of simulated PSD D_j^{rep}. Thus, the quantity $D_{ij}^{rep} (=\{D_{1j}^{rep}, \ldots, D_{Kj}^{rep}\})$ can be written as

$$D_{ij}^{rep} = \frac{B_{ij}X_{ij}}{2}. \tag{9}$$

We can then get K values of T_R^{rep} (see Equation (8)) to calculate the distribution of T_R^{rep}.

At the same time, we get the best-fit model M_j of the current model through the Bayesian posterior distribution. After calculating the distribution of T_R^{rep}, we take the PSD of the input signal as D_j^{obs} to calculate T_R^{obs}. We can then calculate the Bayesian p-value as

$$p = \int_{T_R^{obs}}^{+\infty} f(T_R^{rep})dT_R^{rep} = Pr(T_R^{rep} \geqslant T_R^{obs}). \tag{10}$$

The Bayesian p-value is the proportion of the tail area from T_R^{obs} in the probability density distribution of the statistic T_R^{rep} (see Figure 1(d)). A relatively small Bayesian p-value indicates that the current model cannot well interpret the PSD of the input signal. In this step, we disregard a model with $p < 0.01$ (Inglis et al. 2016). On the contrary, the measured test statistic value T_R^{obs} is very small compared to the T_R^{rep} obtained from Bayesian posterior distribution, giving a Bayesian p-value very close to 1, which would be an indication of over-fitting. Thus, we discard a model if $p > 0.99$ too.

3. Evaluating the Performance of Bayesian-based MCMC

In this section, we evaluate the accuracy of the QPP detection method described in Section 2 by applying it to 100 synthetic signals. The signals consist of a superposition of power law and white noise, and an oscillatory signal which mimics a QPP. In different synthetic signals, relative amplitudes of the oscillatory component are randomly generated in the range [0, 0.15] and their instantaneous frequencies are also random values with a normal distribution, 0.004 ± 0.001 Hz in 50 synthetic signals, and 0.008 ± 0.001 Hz in 50 other signals. Figure 1(a) shows a typical synthetic signal in the time domain with a built-in QPP with a mean frequency of
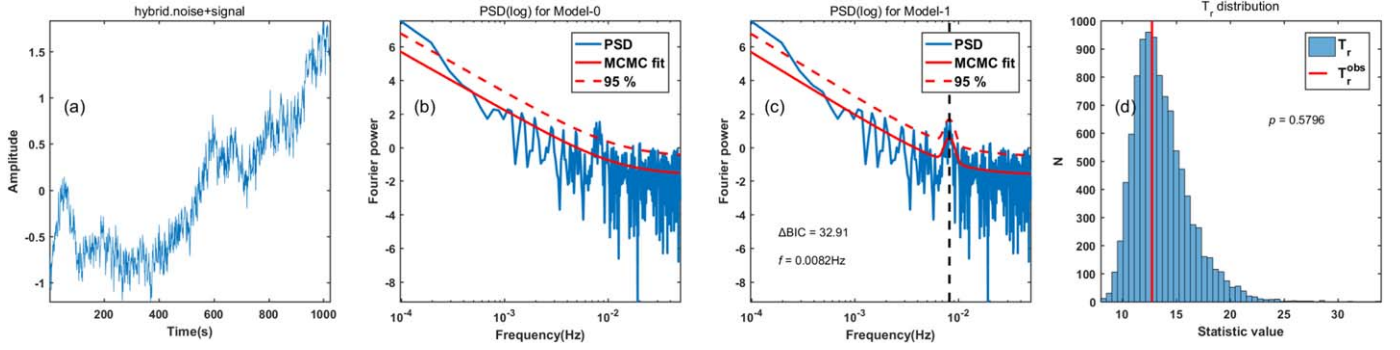
Figure 1. An example of analysis of a synthetic signal. (a) The time series of a synthetic signal which consists of colored and white noises and a short-living oscillation train. (b) The PSD and its fitting with model M_0. The solid-red curve is the model composed of the best parameters obtained by MCMC sampling. The red-dashed curve is its 95% confidence level. (c) The PSD fitted by model M_1. The black-dashed line indicates the central frequency of the best-fitting Gaussian, 0.0082 Hz. In this example, ΔBIC is equal to 32.91, meaning that model M_1 is more suitable than model M_0. (d) The distributions of T_R^{rep} and T_R^{obs} from the best-fitting model M_1. The value $p = 0.5796$ indicates that the PSD can be well interpreted by model M_1.
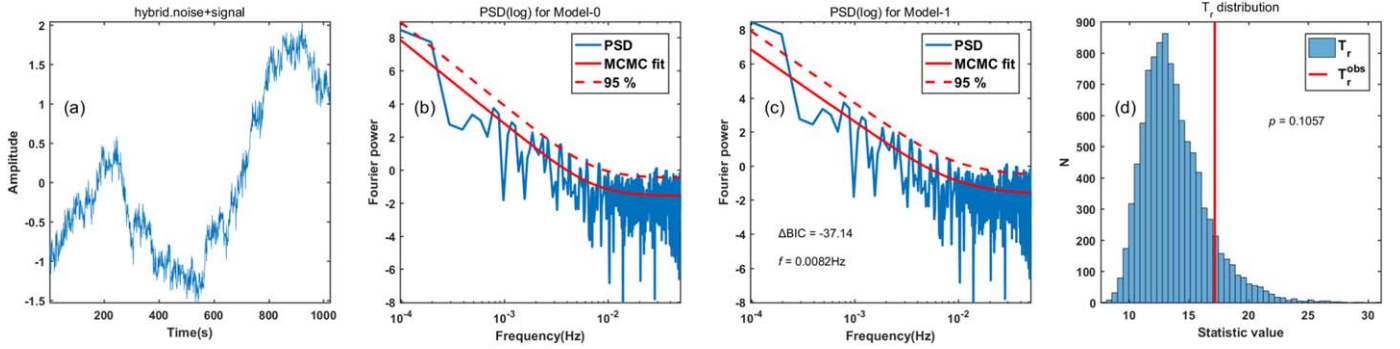

Figure 2. An example of analysis of a synthetic signal that contains only the colored and white noises. From left-hand to right-hand, the panels show the synthetic signal, the M_0 model fitting to the PSD, the M_1 model fitting, and the distribution of T_R^{rep} and T_R^{obs}, respectively.

0.008 Hz. The PSD and its best-fitting with model M_0 (see Equation (4)) are shown in Figure 1(b). The solid-red line is the best-fit model obtained by Bayesian-based MCMC and the red-dashed line is its 95% confidence level. We can see that the peak with a certain width near frequency 0.0082 Hz exceeds the 95% confidence level in Figure 1(b). This means that the peak is unlikely to come from colored or white noise, and hence can be interpreted as a QPP.

Figure 1(c) shows the best-fitting result of M_1 (see Equation (5)). The black-dashed line marks the Gaussian peak position whose frequency value is 0.0082 Hz. The ΔBIC (see Equation (7)) value is then calculated to compare models M_0 and M_1. The value of ΔBIC is equal to 32.91, which means that model M_1 better suits the given PSD than model M_0. Figure 1(d) presents the distributions of T_R^{rep} and T_R^{obs} from model M_1, The Bayesian p-value described in Section 2.4 is $p = 0.5796$, which indicates that the PSD can be well interpreted by this model (i.e., model M_1).

We also test the QPP detection technique on some synthetic signals that contain only the colored and white noise. In Figure 2, we use the same processing for a synthetic signal that contains only the colored and white noise. Figure 2(c) shows there is no obvious peak to fit in the PSD and the value of ΔBIC is negative, which means that model M_0 better suits the given PSD than model M_1. Figure 2(d) presents the distributions of T_R^{rep} and T_R^{obs} from model M_0. Our experiments demonstrate that this technique does not detect false peaks in purely noisy environments and allows for the confident detection of a built-in oscillatory pattern in noisy data.
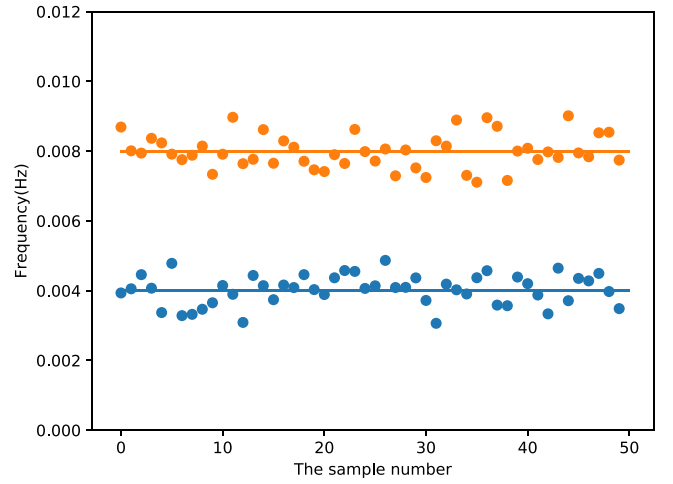

Figure 3. Frequencies of oscillatory components determined with use of the Bayesian-based MCMC method in 100 synthetic noisy signals. Half of the synthetic signals have a built-in oscillatory signal with a frequency 0.008 ± 0.001 Hz (the orange-horizontal line), while the half have a frequency of 0.004 ± 0.001 Hz (the blue-horizontal line).

Figure 3 presents the estimations of the oscillation frequency in our 100 synthetic signals, which are made with the use of Bayesian-based MCMC. The orange and blue lines represent the two mean values of the frequencies of two built-in oscillatory components, 0.004 Hz and 0.008 Hz, respectively. A total of 50 orange dots and 50 blue dots mark the positions of the detected frequencies. They fluctuate around the means, 0.004 Hz or 0.008 Hz, which demonstrates that the results are
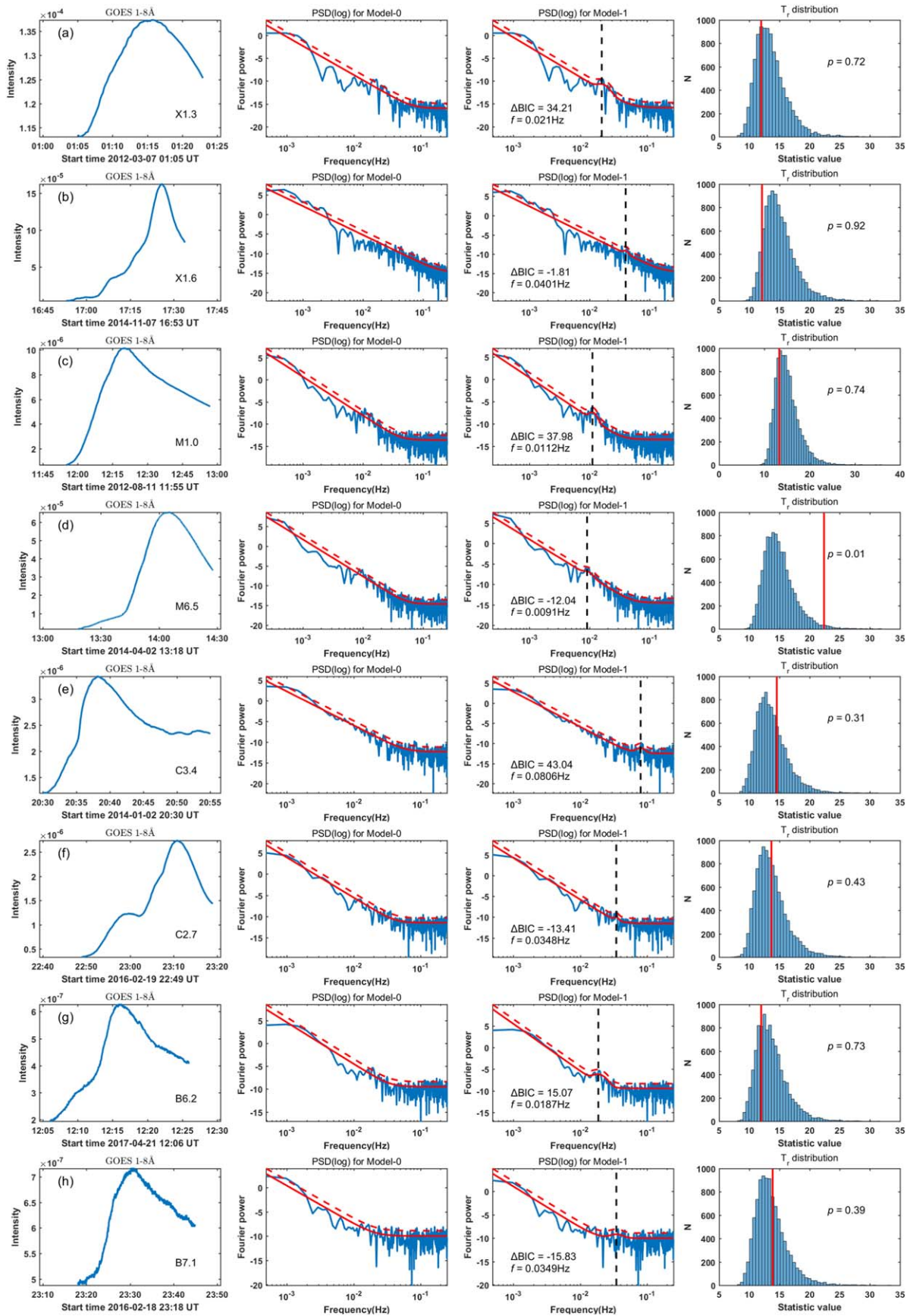
Figure 4. Examples of eight solar flare lightcurves obtained in the soft X-ray band. In rows, the panels show the original signal, the M_0 model fitting to the PSD, the M_1 model fitting, and the distribution of $T_R^{\rm rep}$ and $T_R^{\rm obs}$, respectively.
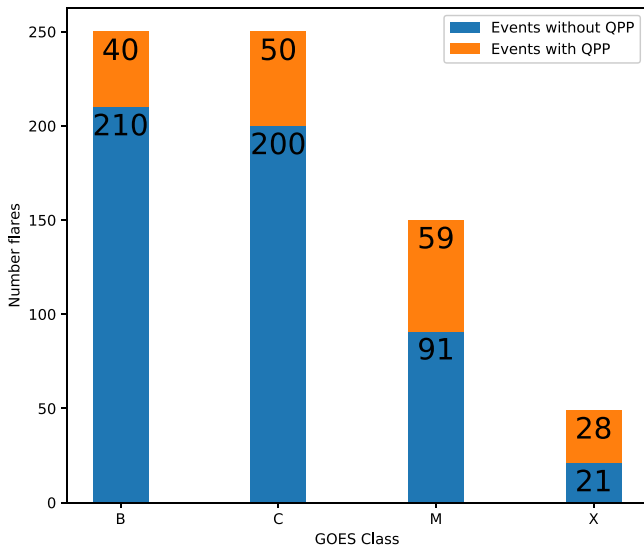
Figure 5. The percentage of flares with detected QPP. There are a total of 28 QPP events detected in 49 X-class flares, 59 in 150 M-class flares, 50 in 250 C-class flares, and 40 in 250 B-class flares.

very close to the given frequencies. The distribution results shown in Figure 3 indicate that the methodology described in Section 2 is accurate and feasible.

4. Analysis of GOES Flare Signals

A total of 699 flare events observed at soft X-rays with GOES were randomly selected from 2010 to 2017, including 250 B-class flares, 250 C-class flares, and 150 M-class flares, as well as all 49 X-class flares. The flare signals were recorded by the XRS instrument on GOES-15 in the 1–8 Å channel. The start and end times of the time series are selected according to the GOES flare catalog.

Figure 4 shows analysis of eight typical flare signals, including two samples from each flare class. One of the flares of each class has a detected QPP event and the other is not. For example, Figures 4(a) and (b) show X-class flares. A QPP is detected in the first flare and is not detected in the second. Similar to Figure 1, each row in Figure 4 from left-hand to right-hand shows the time series signal, i.e., the lightcurve, fitting of its PSD with models M_0 and M_1, and the T_R^{rep} and T_R^{obs} distribution.

The processing is the same as that of the synthetic signals shown in Figure 1. In particular, Figure 4(a) shows an X1.3 flare which occurred on 2012 March 7 and lasted for about 18 minutes. We used models M_0 and M_1 to fit its PSD. We then determined the best-fitting parameters of each model and calculated their BIC values separately to get ΔBIC. A value of ΔBIC = 34.21 indicates that model M_1 is better than model M_0. Thus, the flare has a QPP pattern which is manifested by a Gaussian peak centered at 0.021 Hz (i.e., the period about 47.62 s). The Bayesian p-value for Model 2 is $p = 0.72$ m, which indicates that the PSD of the input signal can be well interpreted by the current model. In summary, the results demonstrate that there is a quasi-periodic signal in this flare and its period is about 48 s.

The solar flares shown in Figures 4(a), (c), (e), and (g) present strong evidence that models M_1 fit well their PSDs because ΔBIC > 10 and $0.99 \geqslant p \geqslant 0.01$. Thus, we conclude that the QPP is presented in these four flares. On the contrary,
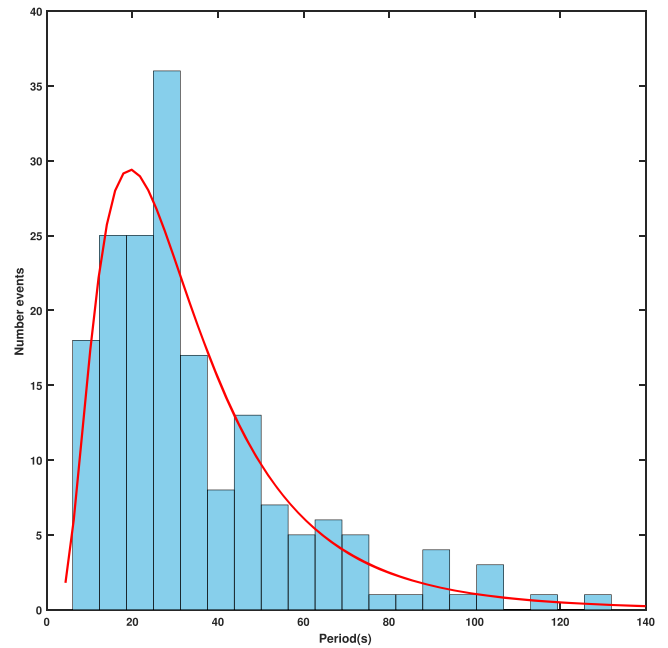

Figure 6. The period distribution in the detected QPP. We detected 177 QPP events in the 699 flares. The periods are mainly in a range between 6.2 and 75.3 s. The minimum and maximum periods are 6.3 s and 130.2 s, respectively.

Figures 4(d), (f), and (h) strongly favor model M_0 because ΔBIC < −10 and $0.99 \geqslant p \geqslant 0.01$, which means that the QPP is not detected in these three flares. The analysis of the flare shown in Figure 4(b) gives us insufficient evidence ($|\Delta$BIC$| < 10$), and hence we are unable to discriminate between these two models.

In addition, the flare shown in Figure 4(d) has an additional peak in the high frequency part of the spectrum, at 0.1827 Hz (the period about 5.47 s). Because the sampling interval of GOES-15 satellite is about 2 s, according to the Nyquist sampling theorem we cannot determine whether or not this peak is caused by a quasi-periodic signal.

The Bayesian-based MCMC method led to the QPP detection in 177 out of 699 analyzed flares, in other words, approximately in 25.5% of the analyzed events. Figure 5 illustrates the number of the QPP detections in flares of various classes. The orange part of each bar represents the proportion of QPP events and the blue part represents the proportion of events without QPP. A total of 28 QPP events in 49 X-class flares, 59 in 150 M-class flares, 50 in 250 C-class flares, and 40 in 250 B-class flares are determined by the Bayesian-based MCMC method. Approximately 43.7% of X- and M-class GOES events show strong evidence of QPP, which is higher than the result (∼30%) obtained in Inglis et al. (2016). The fractions of flares with detected QPP in X-, M-, and C-class GOES events are higher than the results of Hayes et al. (2020). In addition, our results indicate that QPPs are detected more frequently in flares of more powerful classes.

Figures 6 and 7 present the distributions of the detected QPP periods. In Figure 6, we show the distribution of all periods, of 177 QPP detected in 699 flares. A total of 165 of the 177 QPP are found to have periods between 6.2 and 75.3 s, and 121 of the 177 QPP events have periods between 6.2 and 37.5 s. The minimum period is 6.28 s and the maximum period is 130.2 s. The distribution is close to the logarithmic normal distribution, which is consistent with the results of Hayes et al. (2020).
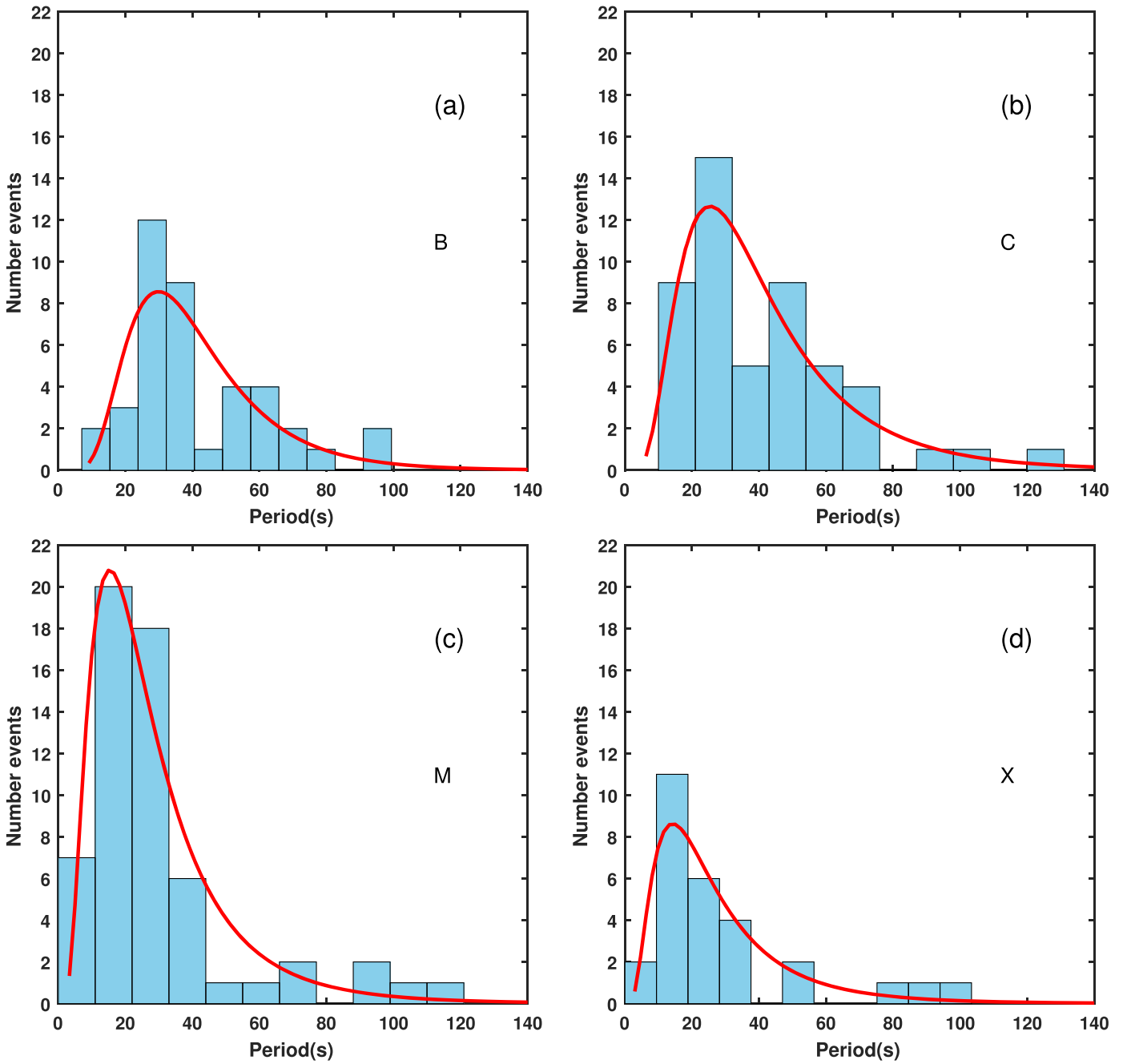
Figure 7. Distributions of QPP periods in flares of B-, C-, M-, and X- classes are shown panels (a)–(d), respectively. In panel (a), there are 40 QPP periods detected in 250 B-class flares, 38 of the 40 periods are between 12.2 and 74.6 s. In panel (b), 47 of the 50 periods detected in 250 C-class flares range between 11.5 and 74.9 s. In panel (c), 52 of the 59 periods detected in M-class flares are found to range between 7.5 and 48.0 s, and in panel (d), in X-class flares, 22 of the 28 periods are between 6.2 and 32.0 s.

In Figure 7, the distribution of the QPP periods in B-, C-, M-, and X-class flares are shown in panels (a)–(d), respectively. In panel (a), there are 40 periods of QPP events from 250 B-class flares, 38 of the 40 periods are between 12.2 and 74.6 s. In panel (b), 47 of the 50 periods from 250 C-class flares are between 11.5 and 74.9 s. In panel (c), 52 of the 59 periods are between 7.5 and 48.0 s. In panel (d), 22 of the 28 periods are between 6.2 and 32.0 s. The distributions of QPP periods in different flare classes are similar, which indicates that the specific value of a QPP period does not depend on the flare's power.

5. Discussion and Conclusion

We performed a search for stationary QPP patterns in the thermal emissions generated in solar flares. The QPP detection was carried out with the use of the Bayesian-based MCMC technique. The Fourier power spectra (PSD) of the time signals (i.e., the lightcurves) were fitted with two candidate models: a power-law-plus-constant model M_0 and a power-law-plus-constant model with an additional spectral peak of the Gaussian shape M_1. The latter model, M_1, represents a typical spectrum of a flare with a QPP pattern (e.g., Pugh et al. 2017a; Broomhall et al. 2019). The models were compared with the

use of BIC, which takes into account the maximum likelihood, the number of model parameters, and the number of data points. This approach punishes the model with more parameters more heavily, which is conducive to improving the accuracy and reliability of the experimental results. In addition, we calculated the Bayesian p-value. When the Bayesian p-value is within the range of 0.01...0.99, we conclude that this best-fitting model explains well the PSD of the analyzed lightcurve. If the best-fitting model is model M_1, then we conclude that the analyzed flare has a QPP pattern. The feasibility of this technique was tested on 100 synthetic signals, which consist of the superposition of colored and white noises and an oscillation.

We analyzed 699 solar flares of various classes, which were observed with GOES-15 from 2010 to 2017. We summarize the main findings as follows:

1. Approximately 25.3% (177 cases) of GOES flares show strong evidence of QPP. The proportions of QPP events in the four classes of flares are: $\sim$57% (28) in the X-class, 39% (59) in the M-class, 20% (50) in the C-class, and 16% (40) in the B-class. Thus, QPP is detected more frequently in more powerful flares.

2. A total of 165 of the 177 QPP events have periods between 6.2 and 75.3 s, and 121 of the 177 QPP events have periods between 6.2 and 37.5 s. The minimum period is 6.3 s and the maximum period is 130.2 s. The population distribution is close to the logarithmic normal distribution. In class B, 38 of the 40 QPP periods are between 12.2 and 74.6 s. In class C, 47 of the 50 QPP periods are between 11.5 and 74.9 s. In class M, 52 of the 59 QPP periods are between 7.5 and 48.0 s. In class X, 22 of the 28 QPP periods are between 6.2 and 32.0 s. The distributions of QPP periods in each class of flares are similar. We conclude that a QPP period is independent of the flare power. This may indicate that the parameters of QPP are determined by geometric parameters of the flaring active region (e.g., the characteristic length and minor radius of coronal loops) rather than the speed and power of the magnetic reconnection process. This suggests that QPP could be connected with magneto-hydrodynamic oscillatory processes in the flaring active region (e.g., McLaughlin et al. 2018; Zimovets et al. 2021), the parameters of which are determined by the geometric sizes of plasma non-uniformities.

3. In general, QPP is the result of a variety of influence factors (e.g., McLaughlin et al. 2018; Zimovets et al. 2021). The population distribution of QPP periods is close to the logarithmic normal distribution rather than the normal distribution, which indicates that these factors are not independent events. Consequently, we speculate that the influence factors of QPP are non-independent.

The performed analysis may have a few limitations. First, in PSD, if the power of QPP is low, then it will be drowned by background noise and so we cannot detect such QPP. Data acquisition may also be affected by the sampling environment and the instrument itself. Second, because we pursue the accuracy and reliability of the estimation, the experimental process is relatively rigorous at each step. Take the width σ of the Gaussian component as an example. The instantaneous frequency of QPP in a flare signal could change with time, which widens the corresponding spectral peak. In our work, to ensure that this obvious feature of QPP can be extracted, we made certain restrictions on the width σ of the Gaussian component of model M_1. This may lead to a situation where model M_1 cannot completely capture this feature of QPP, which leads to a lack of detection. Likewise, a QPP pattern may be localized in time (i.e., be an oscillation train). This non-stationary feature could also decrease the detection rate. Another complication could be connected with oscillatory patterns of a non-harmonic shape. In this case, the spectral power of the oscillation is distributed over several frequencies, which reduces the power of the main spectral peak. In addition, the flares that we randomly selected were of relatively long duration, and thus the lightcurves had a sufficiently large number of data points for a Fourier transform. At the same time, to avoid making the data fitting time too long, we did not choose flares with a very long duration, so the durations of the four types of flares in the experiment are basically within the same range. We speculate that the occurrence rate of QPP events is independent of flare duration. Thus, we should admit that our QPP detection procedure could miss QPP of certain types. This means that the QPP appearance rate that is determined in this study is rather a lower boundary.

A series of methods used in these experiments can also be applied to the detection of QPP in fixed flare signal and other application scenarios that need to be sampled in high-dimensional space.

ORCID iDs

Bo Liang ● https://orcid.org/0000-0002-0604-0949
Song Feng ● https://orcid.org/0000-0003-4709-7818
Ding Yuan ● https://orcid.org/0000-0002-9514-6402
Valery M. Nakariakov ● https://orcid.org/0000-0001-6423-8286
Yunfei Yang ● https://orcid.org/0000-0001-9927-7541

References

Abramov-Maximov, V. E., & Bakunina, I. A. 2020, Ge&Ae, 59, 822
Altyntsev, A. T., Meshalkina, N. S., Sych, R. A., & Kolotkov, D. Y. 2022, A&A, 663, A149
Anfinogentov, S. A., Antolin, P., Inglis, A. R., et al. 2022, SSRv, 218, 9
Anfinogentov, S. A., Nakariakov, V. M., Pascoe, D. J., & Goddard, C. R. 2021, ApJS, 252, 11
Broomhall, A.-M., Davenport, J. R. A., Hayes, L. A., et al. 2019, ApJS, 244, 44
Burnham, K. P., & Anderson, D. R. 2004, Sociological Methods & Research, 33, 261
Chatfield, C. 2003, The Analysis of Time Series: An Introduction (London: Taylor and Francis)
Clarke, B. P., Hayes, L. A., Gallagher, P. T., Maloney, S. A., & Carley, E. P. 2021, ApJ, 910, 123
Hastings, W. K. 1970, Biometrika, 57, 97
Hayes, L. A., Gallagher, P. T., Dennis, B. R., et al. 2019, ApJ, 875, 33
Hayes, L. A., Inglis, A. R., Christe, S., Dennis, B., & Gallagher, P. T. 2020, ApJ, 895, 50
Inglis, A. R., Ireland, J., Dennis, B. R., Hayes, L., & Gallagher, P. 2016, ApJ, 833, 284
Inglis, A. R., Ireland, J., & Dominique, M. 2015, ApJ, 798, 108

Kolotkov, D. Y., Nakariakov, V. M., Kupriyanova, E. G., Ratcliffe, H., & Shibasaki, K. 2015, A&A, 574, A53

Kolotkov, D. Y., Pugh, C. E., Broomhall, A.-M., & Nakariakov, V. M. 2018, ApJL, 858, L3

Kupriyanova, E. G., Melnikov, V. F., Nakariakov, V. M., & Shibasaki, K. 2010, SoPh, 267, 329

Li, D., Feng, S., Su, W., & Huang, Y. 2020a, A&A, 639, L5

Li, D., Li, Y., Lu, L., et al. 2020b, ApJL, 893, L17

Liang, B., Meng, Y., Feng, S., & Yang, Y. 2020, Ap&SS, 365, 40

McLaughlin, J. A., Nakariakov, V. M., Dominique, M., Jelínek, P., & Takasao, S. 2018, SSRv, 214, 45

Milligan, R. O., Fleck, B., Ireland, J., Fletcher, L., & Dennis, B. R. 2017, ApJL, 848, L8

Nakariakov, V. M. 2007, AdSpR, 39, 1804

Nakariakov, V. M., Anfinogentov, S., Storozhenko, A. A., et al. 2018, ApJ, 859, 154

Nakariakov, V. M., Foullon, C., Myagkova, I. N., & Inglis, A. R. 2010, ApJL, 708, L47

Nakariakov, V. M., Kolotkov, D. Y., Kupriyanova, E. G., et al. 2019, PPCF, 61, 014024

Nakariakov, V. M., & Melnikov, V. F. 2009, SSRv, 149, 119

Pugh, C. E., Broomhall, A. M., & Nakariakov, V. M. 2017a, A&A, 602, A47

Pugh, C. E., Nakariakov, V. M., Broomhall, A. M., Bogomolov, A. V., & Myagkova, I. N. 2017b, A&A, 608, A101

Schwarz, G. 1978, AnSta, 6, 461

Sharma, S. 2017, ARA&A, 55, 213

Simões, P. J. A., Hudson, H. S., & Fletcher, L. 2015, SoPh, 290, 3625

Tan, B., & Tan, C. 2012, ApJ, 749, 28

Tan, B., Yu, Z., Huang, J., Tan, C., & Zhang, Y. 2016, ApJ, 833, 206

Tian, H., Young, P. R., Reeves, K. K., et al. 2016, ApJL, 823, L16

Vaughan, S. 2005, A&A, 431, 391

Vaughan, S. 2010, MNRAS, 402, 307

Yuan, D., Feng, S., Li, D., Ning, Z., & Tan, B. 2019, ApJL, 886, L25

Zimovets, I. V., McLaughlin, J. A., Srivastava, A. K., et al. 2021, SSRv, 217, 66

Zimovets, I. V., Nechaeva, A. B., Sharykin, I. N., & Nizamov, B. A. 2022, Ge&Ae, 62, 356

Zimovets, I. V., & Struminsky, A. B. 2009, SoPh, 258, 69