RAS Techniques and Instruments

RASTAI **2,** 238–255 (2023) https://doi.org/10.1093/rasti/rzad015
Advance Access publication 2023 May 18

Searching for quasi-periodic eruptions using machine learning

Robbie Webbe ★ and A. J. Young
H. H. Wills Physics Laboratory, Tyndall Avenue, Bristol BS8 1TL, UK

Accepted 2023 May 5. Received 2023 May 2; in original form 2023 February 2

ABSTRACT

Quasi-periodic eruption (QPE) is a rare phenomenon in which the X-ray emission from the nuclei of galaxies shows a series of large amplitude flares. Only a handful of QPEs have been observed but the possibility remains that there are as yet undetected sources in archival data. Given the volume of data available a manual search is not feasible, and so we consider an application of machine learning to archival data to determine whether a set of time-domain features can be used to identify further light curves containing eruptions. Using a neural network and 14 variability measures we are able to classify light curves with accuracies of greater than 94 per cent with simulated data and greater than 98 per cent with observational data on a sample consisting of 12 light curves with QPEs and 52 light curves without QPEs. An analysis of 83 531 X-ray detections from the XMM Serendipitous Source Catalogue allowed us to recover light curves of known QPE sources and examples of several categories of variable stellar objects.

Key words: machine learning – galaxies: nuclei – X-rays: galaxies.

1 INTRODUCTION

The role of machine learning in astrophysics is becoming progressively important, with the scope and scale of surveys and planned missions resulting in ever-increasing volumes of data and in growing archives for current missions. Citizen science projects, like Galaxy Zoo (e.g. Lintott et al. 2008) will struggle to cope with the volume of data that is expected to be produced with planned survey missions. There is a pressing need to develop automated tools which can process and reduce volumes of data to manageable amounts. Transient events of many types have been the focus of several machine learning approaches. Due to their fleeting nature and the time sensitive nature of follow-up efforts, automation has the potential to increase the number of transient events detected, and to allow for them to be detected sooner. This will allow a greater proportion of their lifetimes to be monitored and including automation in processing pipelines also allows for such events to be detected before a scientist could interact with the observed data. Approaches in using supervised and unsupervised machine learning to detect and classify a greater proportion of transient events in (near) real time (e.g. Narayan et al. 2018; Muthukrishna et al. 2019a, 2022; Muthukrishna, Parkinson & Tucker 2019b) using optical observations have allowed supernovae and some other classes of transients to be detected during the course of the events, although understandably the accuracy of these techniques increases as more of the events are detected. High-energy data present different challenges, as the statistics underpinning observed data are different due to the typically low count rates. Attempts at detecting X-ray transient sources using supervised learning (random forest methods) with the 2XMM and 3XMM Serendipitous Source Catalogues have achieved accuracies of $\simeq$ 97 per cent (Lo et al. 2014) and $\simeq$ 92 per cent (Farrell, Murphy & Lo 2015) across multiclass classifications using combinations of time-domain and spectroscopic features.

With new classes of X-ray transients like quasi-periodic eruptions (QPEs) continuing to be discovered it is important to develop methods for detecting these new classes both in archival data and continuing and planned surveys as soon as possible to develop our understanding of these transients. If it is possible to detect QPEs with established machine learning techniques and variability features used to detect other types of variability, this could significantly increase our known QPE host population.

The first QPEs were detected in the active galactic nucleus (AGN) GSN 069 by Miniutti et al. (2019), and four further likely sources were subsequently identified in the extragalactic sources RX J1301.9+2747 (Giustini, Miniutti & Saxton 2020), eRASSU J023147.2−102010 (eRO-QPE1), and eRASSU J023448.9−441931 (eRO-QPE2; Arcodia et al. 2021), and XMMSL1 J024916.6−041244 (Chakraborty et al. 2021). Although all AGN show X-ray variability, QPEs are characterized by short lived, large scale changes in X-ray luminosity with eruptions appearing greater in amplitude, peaking at earlier times, and having shorter durations with increasing photon energy bands. Of those objects, two were detected by direct analysis of observations of similar sources (GSN 069 and RX J1301.9+2747, Miniutti et al. 2019; Giustini et al. 2020). The sources eRO-QPE1 and eRO-QPE2 were detected by means of a blind search through data released by the eROSITA instrument, with a simple cut by count rate and significant variability being used to identify sources for further examination (Arcodia et al. 2021). The search which identified the QPE candidate source XMMSL1 J024916.6−041244 used the quasi-periodic automated transit search algorithm (Carter & Agol 2013; Chakraborty et al. 2021) which was originally intended for the identification of exoplanet transits. QPEs appear to be transient phenomena, with the eruptions seen in XMMSL1 J024916.6−041244 not appearing in an observation 15 yr after that in which they were observed, and the

★ E-mail: robbie.webbe@bristol.ac.uk

eruptions in GSN 069 having not been observed in the most recent observations (Miniutti et al. 2023) following an extended period when they were visible.

The mechanism which causes the appearance of QPEs is currently unknown, but possible explanations which have been proposed for the phenomena include: accretion from an orbiting object (King 2020, 2022; Chen et al. 2022; Krolik & Linial 2022; Lu & Quataert 2022; Wang et al. 2022; Zhao et al. 2022; Linial & Sari 2023); collision between an orbiting body and the AGN's accretion disc (Suková et al. 2021; Xian et al. 2021; Franchini et al. 2023; Linial & Metzger 2023); tearing of warped accretion discs (Raj & Nixon 2021; Musoke et al. 2023); accretion from interacting stellar extreme mass ratio inspirals (Metzger, Stone & Gilbaum 2022); disc instabilities (Sniegowska et al. 2020; Kaur, Stone & Gilbaum 2022; Pan et al. 2022); and gravitational lensing in supermassive black hole binary systems (Ingram et al. 2021). Ultimately, in order to provide a greater evidence base upon which to make rigorous determinations as to the true mechanism for the creation of QPEs we need to identify more sources, and so exploiting future surveys as well as archival data will be important.

In this paper, we aim to determine whether light curves containing QPEs can be distinguished from those which do not by means of an automated system based upon a series of time-domain variability features. For the purposes of this analysis we will be focusing on the temporal features of QPEs as seen in the five host objects identified to date. We aim to identify patterns of variability which are quasi-Gaussian in appearance with long periods of quiescence between eruptions. This will allow us to identify future QPE candidates in large survey data or by searching through archival databases. We do this by means of a neural network trained on simulated data and also test the resulting classifier on real observational data from *XMM–Newton*. In Section 2, we describe the generation of the simulated training data sets, the variability features to be used and how the optimal architecture for the classifiers will be determined. We report the results of the classifier against the simulated and real data sets in Section 3. In Section 4, we discuss the results of the classifier and how it can be used to exploit readily available archival data catalogues, and in Section 5, we consider the performance of the classifier and avenues for future work.

2 METHODS

2.1 Observational data preparation

The training data which we have used in this analysis are derived from a series of *XMM–Newton* observations, listed in Table 1, of the QPE sources GSN 069, RX J1301.9+2747, XMMSL1 J024916.6−041244, eRASSU J023147.2−102010, and eRASSU J023448.9−441931. We obtained the data from the XMM Science archive[1] and reprocessed all observations using XMM Science Analysis System.[2] For all light curves, the EPIC pn camera event data were extracted for photon energies in the range 0.2–2.0 keV, background and barycentre corrected, and was binned at a rate of 10 s. We use photon energies in this range as eruptions in this energy range have been easily detectable in known QPE sources and have previously been used to characterize QPE profiles (Miniutti et al. 2019; Giustini et al. 2020, etc.). The resulting light curves were then manually screened for flaring events before being passed for analysis.

[1] http://nxsa.esac.esa.int/nxsa-web/
[2] version xmmsas_20190531_1155-18.0.0.

Table 1. Details of *XMM–Newton* observations of five QPE host galaxies used in generating the training data. Observation exposures and numbers of eruptions contained are listed in the last two columns.

Object	OBSID	Exposure (ks)	Erup.
GSN 069	0823 680 101	63.3	2
–	0831 790 701	141.4	5
–	0851 180 401	135.4	5
–	0864 330 101	141.0	4
RX J1301.9+2747	0124 710 801	29.8	1
–	0851 180 501	48.4	3
–	0864 560 101	134.9	8
XMMSL1 J024916.6−041244	0411 980 401	11.7	1
eRASSU J023147.2−102010	0861 910 201	94.2	2
–	0861 910 301	90.2	1
eRASSU J023448.9−441931	0872 390 101	95.0	9
–	0893 810 501	25.0	3

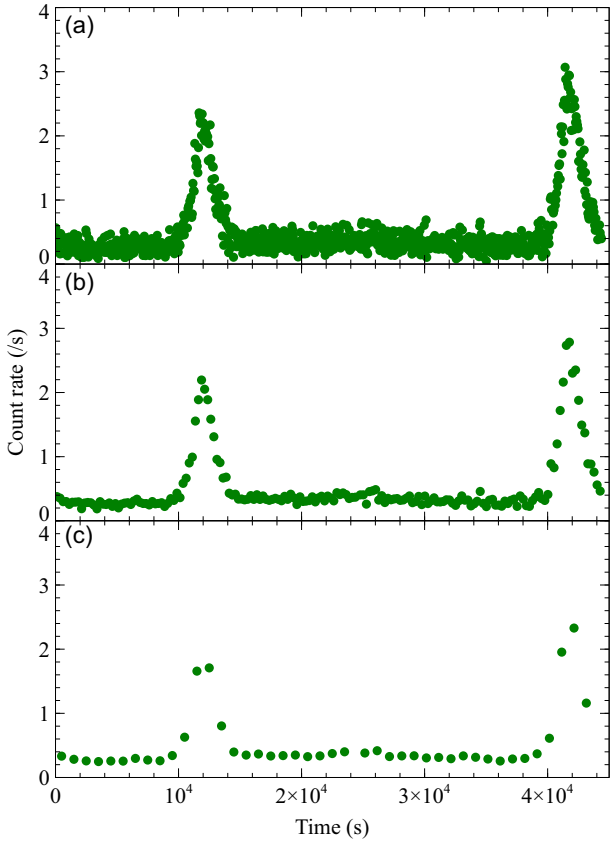

Figure 1. Example of an X-ray light curve containing QPEs. Light curve is from *XMM–Newton* observation 0823680101 of GSN 069, where the counts are binned at rates of (a) 50 s, (b) 250 s, and (c) 1000 s.

An example of a screened light curve containing QPEs is displayed in Fig. 1.

For each of these observations we then created light-curve segments which isolated each eruption, giving a total of 43 segments, each containing one eruption, from 12 observations, and fit a model to the segments of the form

$$x(t) = x_{\rm q} + A \exp\left(\frac{-\ln(2)(t - t_{\rm peak})^2}{t_{\rm dur}^2}\right) \quad (1)$$

where x_q is the quiescent baseline count rate in the vicinity of each eruption, A is the amplitude of the eruption relative to this quiescent count rate, t_{peak} is the time of the peak of the eruption within the segment, and t_{dur} is the duration of the eruption as defined as the full width at half-maximum of the Gaussian profile of the eruption. For the purposes of the feature extraction we only include the first eruption seen in observation 0861910201 of eRASSU J023147.2−102010, as the second eruption appears to overlap with the first, and there is no 'quiescent' phase on either side of the peak as the observation ends before the eruption has completed. As such there is no baseline against which the peak height could be compared for this second eruption and was therefore not suitable for feature extraction. We accept that excluding the features of this eruption could negatively impact the ability of the classifier to detect other sources with overlapping eruptions. We then used the positions of the peaks within the eight observations for which there were multiple profiled eruptions to determine the recurrence times between eruptions for each observation, and thus the average duty cycle (Δ) for each observation. The duty cycle is calculated as

$$\Delta = \frac{t_{dur}}{t_{rec}}. \tag{2}$$

For the purposes of testing the neural networks against real observational data, we also include 57 light curves from other low-mass AGN. These light curves are of AGN which were selected by Webbe & Young (2023) as part of a targeted search for QPE host galaxies selected by mass from the *Chandra* ACIS Survey for X-ray AGN in Nearby Galaxies catalogue (She, Ho & Feng 2017) which showed no signs of QPE-like behaviour. Details of all observations used in the final evaluation phase of the neural network performance can be found in Table A1 in Appendix A.

2.2 Simulated light-curve generation

In order to produce a training data set of an appropriate size for the training of the neural network, we used the features of the eruptions already observed (amplitude, duration, and duty cycle) and the algorithm for generating light curves outlined in Timmer et al. (1995). We generate light curves with total durations of 100 ks, and time bins of 50, 250, and 1000 s. We use a value of 250 s for the time binning as it fits typical binning values used in the literature, and then the choices of 50 and 1000 s were made as values which could show the variability at the lower level which might be expected to be affected by Poisson noise, and a higher level at which we start to reach the durations of individual eruptive events. With these time bins and considering the typical count rates seen in the data we find that including Poisson noise is not required when training the model (see Section 5.1). A simple power-law model was chosen as a typical underlying PSD of the form $f^{-\beta}$ with values of β being randomly drawn from a normal distribution with $\overline{\beta} = 2.06$ and $\sigma_\beta = 0.01$ as per González-Martín & Vaughan (2012). Values for the index of the power spectra are generated with NUMPY. As the generated light curves have $\bar{x} = 0$, we shift the resulting light curves up by $(10 + \delta)x_{min}$, where δ is sampled from a normal distribution, and x_{min} is the lowest point in the raw light curve. In order to create a sub-population of light curves which contain QPE signals we then multiply half of the simulated light curves with a signal of the form

$$x(t) = 1 + A\sum_{m=1}^{M} \exp\left(\frac{-\ln(2)(t - t_0 - [m-1]t_{rec})^2}{t_{dur}^2}\right), \tag{3}$$

where A is the amplitude of the eruptions being modelled, t_{dur} is the duration, t_0 is the peak time of the first eruption in the series, t_{rec} is the recurrence time between eruption peaks, and where M satisfies

$$M = \left\lfloor \frac{T}{t_{rec}} \right\rfloor + 3. \tag{4}$$

This gives us a balanced data set for training and validation. To simulate a range of possible eruption profiles, we randomly sample values for A, t_0, t_{dur}, and Δ on the basis of the eruptions profiled in Section 2.1. Values for t_0 are sampled from a uniform distribution such that $t_0 \in [-t_{rec}, 0)$, and the values for A, t_{dur} and Δ are sampled from exponentially modified Gaussian distributions (e.g. Gladney, Dowden & Swalen 1969; Grushka 1972). We sample from such a distribution as it provided a high-quality fit to the distribution of observed parameter values. These simulated light curves are then processed to create non-parametric statistical features, as outlined in Section 2.3.

2.3 Variability measures

We will be using a set of variability statistics calculated from our simulated sample, and then from the real testing sample, to determine whether a light curve does or does not contain QPEs. Some of these features have been used previously in attempts to characterize light-curve variability (Sokolovsky et al. 2017), and some have been used in coordination with machine learning techniques like self-organizing maps (Faisst et al. 2019) to characterize AGN variability. Due to the use of simulated light curves in the training and validation steps of our machine learning workflow, we do not have errors on the simulated light curves. As such, we choose 14 features which do not rely on the existence of uncertainties within the data being analysed. We extract these features from the simulated light curves and the observational data at all three time bins of 50, 250, and 1000 s in order to observe the effects of time binning on detection accuracy.

2.3.1 Feature 1 - standard deviation normalized by the mean

In order to remove any issue caused by the different count rates for observations, we use the standard deviation divided by the mean count rate

$$f_1 = \frac{\sqrt{\frac{\sum_i (x_i - \bar{x})^2}{N-1}}}{\bar{x}}, \tag{5}$$

where N is the number of points in the light curve, x_i is the count rate of individual points on the light curve, and $\bar{x}$ is the average count rate.

2.3.2 Features 2–7 - proportion of the light curve further than 1σ to 6σ from the mean

A light curve where the points are normally distributed about a mean count rate will have proportions of points at given numbers of σ from the mean count rate. These features are calculated as

$$f_n = \frac{N_{x,(n-1)\sigma}}{N}, n \in [2, 7], \tag{6}$$

where $N_{x,\,(n\,-\,1)\sigma}$ is the number of points where $|x_i - \bar{x}| > (n-1)\sigma$.

2.3.3 Feature 8 - inter-quartile range normalized by σ

For a light curve where the points are normally distributed about the mean, the middle 50 per cent of points should be within 0.674σ of

the mean, and so it should be that $IQR = 0.674\sigma$.

$$f_8 = \frac{IQR}{\sigma} \tag{7}$$

2.3.4 Feature 9 - skew

The skewness of a light curve will determine how asymetrically the individual count rates are around the mean value. For this analysis, we use the Fisher–Pearson coefficient of skewness:

$$f_9 = g_1 = \frac{m_3}{m_2^{3/2}}, \tag{8}$$

where

$$m_i = \frac{1}{N}\sum (x - \bar{x})^i \tag{9}$$

and m_i is the bias sampled ith central moment.

2.3.5 Feature 10 - Kurtosis

The Kurtosis of a light curve will determine, relatively, how likely it is that the light curve will contain extreme outlying values. For this analysis, we use the Fisher coefficient of Kurtosis:

$$f_{10} = g_2 = \frac{m_4}{m_2^2} - 3, \tag{10}$$

where m_4 and m_2 are calculated as described in equation (9).

2.3.6 Feature 11 - reverse cross-correlation normalized by σ

To provide a second measure of the asymmetry of the light curves being analysed, we calculate the cross-correlation for the observations with themselves, being reversed along the time axis. To mitigate for different average count rates across observations, we normalize all deviations for individual data points by the standard deviation for the observations as a whole:

$$f_{11} = \sum_i \frac{(x_i - \bar{x})(x_{N-i} - \bar{x})}{\sigma^2}, \tag{11}$$

where x_i are the count rates of individual points in the light curve, x_{N-i} is the counterpart point on the reversed light curve, and N is the number of points in the light curve.

2.3.7 Feature 12 - first maximum of the autocorrelation function

As a measure of coherent periodic variability, we take the height of the first peak of the normalized autocorrelation function. This is the first peak after the autocorrelation function has crossed zero, and is normalized by the autocorrelation at $t = 0$.

$$f_{12} = \max\left(\frac{ACF(t)}{ACF(t=0)}\right) \tag{12}$$

2.3.8 Feature 13 - consecutive same sign deviation proportion

We consider the proportion of sets of three consecutive points from the light curve which all display the same sign deviation from the mean count rate. In some examples in the literature (e.g. Wozniak 2000; Shin, Sekora & Byun 2009), a choice is made to require all three points to be at least some multiple of σ from the mean value, but we choose not to include this extra distinction in order to not inadvertently prejudice our results, or to cause degeneracies with Features 2–7:

$$f_{13} = \frac{N_{3x}}{N-2}, \tag{13}$$

where N_{3x} is the number of sets of three consecutive points in the light curve which have the same sign deviation from $\bar{x}$.

2.3.9 Feature 14 - Von Neumann ratio

The final feature compares the difference in the deviation between successive pairs of points to the total variance of the light curve:

$$f_{14} = \eta = \frac{\sum_{i=1}^{N-1}(x_{i+1} - x_i)^2/(N-1)}{\sigma^2}, \tag{14}$$

where x_i and x_{i+1} are pairs of consecutive points.

2.4 Neural network architecture

The neural networks being applied and evaluated in this work are built using the TENSORFLOW (Abadi et al. 2015) package for PYTHON. We create networks with an input layer, either one or two hidden layers, and an output layer. The input has nodes as required by the number of input features, and the output layer has two nodes to force a binary classification between light curves containing QPEs or not. We use the RELU activation function, and as we are considering a classification problem our loss function of choice is CROSS ENTROPY. Outputs from the network are then passed to a SOFTMAX layer to give the probability that light curves belong to the classes. The hidden layers are allowed to have between 3 and 196 nodes, and the precise final architecture for each input data set is determined by KERAS_TUNER (O'Malley et al. 2019). We use the HYPERBAND tuner from KERAS_TUNER for this optimization. The training and evaluation of networks with the simulated data uses accuracy as the metric to be optimized, which forms part of the determination as to the optimal architecture:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}, \tag{15}$$

where TP is the number of true positive, FP is the number of false positive, TN is the number of true negative, and FN is the number of false negative classifications. For these purposes, we consider those light curves which contain QPEs as being classified as positive, with a probability, as determined by the neural network, of containing QPEs being greater than 50 per cent. When we evaluate the performance against the real data sets we also consider purity, completeness, and the F_1 score for each network where

$$\text{Purity} = \frac{TP}{TP + FP}, \tag{16}$$

$$\text{Completeness} = \frac{TP}{TP + FN}, \tag{17}$$

and

$$F_1\ \text{Score} = \frac{2 \times \text{Purity} \times \text{Completeness}}{\text{Purity} + \text{Completeness}}. \tag{18}$$

We use these additional measures as the observational data set is less balanced, with only 17.4 per cent of the real observations containing QPEs. These metrics are not used to inform further amendments to the architecture of the network, or the weightings applied to individual nodes, but are simply used to illustrate the effectiveness of the network in classifying real observational data.

Table 2. Optimal architecture for the neural networks at each time binning value as determined by `keras_tuner`.

Time bin (s)	No. hidden layers	Nodes
50	1	[110, 2]
250	1	[49, 2]
1000	2	[23, 91, 2]

Table 3. Classification accuracy for 18 000 simulated light curves in the validation phase, and 10 000 in the testing phase.

Time bin (s)	Validation	Testing
50	0.9404	0.9379
250	0.9489	0.9502
1000	0.9467	0.9495

3 RESULTS

3.1 Simulated light-curve classification

We simulated 100 000 light curves, with 50 000 to contain QPEs, and 50 000 without QPEs, with a duration of 100 ks and time binning of 50, 250, and 1 ks in the manner described in Section 2.2 and extracted from them the variability features described in Section 2.3. From the population of 100 000 simulated light curves at each time binning rate, we set aside 10 000 light curves for testing and split the remainder in the ratio 80:20 for training and validation.

The optimal architectures for the neural networks were determined as described in Section 2.4 and allowed to vary between different values for the time binning both for the number of nodes in any hidden layers, and by the number of hidden layers which were allowed to exist. The optimal architectures for the two lower time binned networks included only one hidden layer, while the 1 ks binned network contained two hidden layers. In all cases, there was an output layer which was fixed to contain two nodes in order to force a binary output choice between the light curve containing QPEs or not. The full architectures are given in Table 2.

Across all three time bin values we achieved very high levels of accuracy when classifying the simulated data sets, with the validation and testing accuracy being greater than 94 per cent in all cases and are given in Table 3. Validation and testing accuracy were greatest for the 250 s binned data sets. Differences in performance are, however, only marginal and the slight changes in accuracy upon classifying the simulated testing data set suggests that the model was not significantly overfit to the training data.

3.2 Observational data classification

The observational data, as listed in Table A1, were rebinned to rates of 50, 250, and 1000 s and features were extracted for all available light curves. There were three observations without QPEs which could not be rebinned at a rate of 250 s, and a further two which could not be rebinned at a rate of 1000 s as well due to the limitations on good time intervals. The results of their classification are given in Table 4.

There is no clear hierarchy when classifying real observational data, and at first it appears that all three sets of observational data are classed with equally poor performance, although the 250 s binning performs slightly below the other two classifiers. The

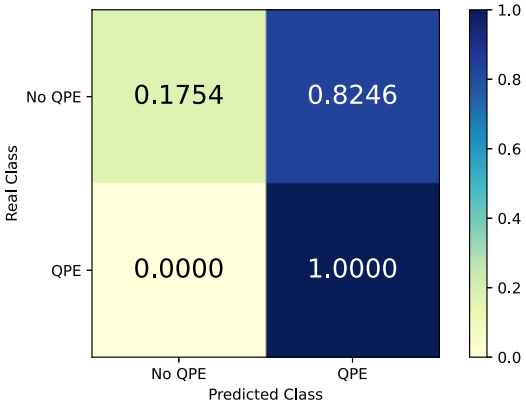

Figure 2. Confusion matrix showing the classification of 57 light curves without QPEs and 12 with light curves with QPEs present. The light curves were binned at a rate of 50 s, and their full details are given in A1. Threshold probability for requiring a QPE is set at 0.500.

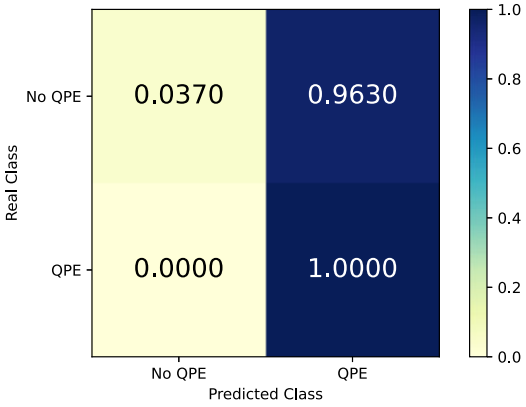

Figure 3. Confusion matrix showing the classification of 54 light curves without QPEs and 12 with light curves with QPEs present. The light curves were binned at a rate of 250 s, and their full details are given in A1. Threshold probability for requiring a QPE is set at 0.500.

completeness of close to 1.0 and low levels of purity suggest that all light curves are being strongly predicted as containing QPEs. The drop in accuracy when classifying real observational data rather than simulated data suggests at first that there may be features of the real observational data which are not appropriately captured in the simulated data sets. The confusion matrices showing the classifications of the observational data against all three classifiers are show in Figs 2, 3, and 4. From the confusion matrices, we can see that the overall accuracy is lowered by large numbers of light curves being incorrectly classified as containing QPEs. When we plot the QPE probabilities of the three classifiers, we can see that the QPE probabilities generally are quite high and close to 1.0 in the majority of cases, as shown in Fig. 5, but that those light curves containing QPEs do tend to be classed more strongly as containing QPEs than those without, and as such we should consider the level at which we delineate between QPE and non-QPE containing light curves. We do consider the benefits of probability calibration, as outlined in Guo et al. (2017), but find that calibration by means of isotonic regression using SKLEARN (Pedregosa et al. 2011) does not beneficially redistribute the probabilities, with the granulation of probabilities close to 1.0 being lost.

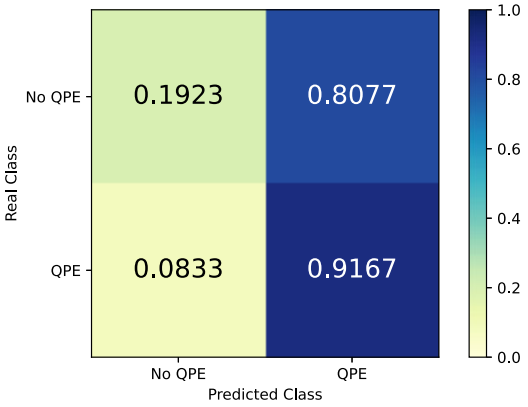

Figure 4. Confusion matrix showing the classification of 52 light curves without QPEs and 12 with light curves with QPEs present. The light curves were binned at a rate of 1000 s, and their full details are given in A1. Threshold probability for requiring a QPE is set at 0.500.

4 DISCUSSION

4.1 Simulated light-curve feature distributions

The simulated light curves were created according to a well-established algorithm which has been used for many purposes (e.g. Hübner et al. 2022; Panagiotou, Kara & Dovciak 2022, etc.), and several of the features selected were chosen to match approaches taken in other attempts to identify variability in time-domain data (e.g. Sokolovsky et al. 2017; Pashchenko, Sokolovsky & Gavras 2018; Faisst et al. 2019, etc.). None of the features were, however, selected with the expectation that they would serve this specific purpose in the identification of light curves containing QPEs. When we look at the features being extracted from the simulated light curves, we can see that there are several all-time binning rates which appear to distinguish well between light curves with and without QPEs. Features 1, 2, 4, 8, 9, and 12 across all-time binning rates are separated by at least 1σ, as shown in Table 5. It should be noted that all of the features except 9 and 11 are constrained to be positive only. In some of the cases the feature distributions cannot be used in all cases to classify the light curve as containing a QPE or not, but specific values may be strong indicators as containing a QPE or not. In the cases of Features 6 and 7, any non-zero value is strongly indicative of a light curve containing QPEs, but not all QPE containing light curves have non-zero values for these features. The values for Feature 1, which appear to be a good way of distinguishing between the two classes, should also be taken with a measure of caution. They are affected by the factor which shifts the light curves from being zero-centred (a necessary step for the imposition of the QPE signals in the light-curve generation), and a further investigation would be needed in order to determine what effect varying this factor has on the distributions for Feature 1 in the two classes and whether it remains a distinguishing factor. If we consider the distributions of the features, as shown in Figs 6, 7, and 8 we can see that for

Table 5. Average and spread on feature values for the 14 features as described in Section 2.3. Values stated are the mean and σ from all 50 000 simulated light curves used in the training, validation, and testing phases for each of the three values for time binning in the samples with and without imposed QPE signals.

Feature	ΔT	QPE sample	Non-QPE sample
f_1	50	0.881 ± 0.449	0.047 ± 0.009
	250	0.883 ± 0.448	0.048 ± 0.010
	1000	0.878 ± 0.449	0.051 ± 0.010
f_2	50	0.134 ± 0.105	0.358 ± 0.047
	250	0.134 ± 0.105	0.358 ± 0.048
	1000	0.135 ± 0.106	0.359 ± 0.051
f_3	50	0.057 ± 0.028	0.029 ± 0.021
	250	0.057 ± 0.028	0.029 ± 0.021
	1000	0.057 ± 0.030	0.029 ± 0.023
f_4	50	0.025 ± 0.018	0.0003 ± 0.0016
	250	0.025 ± 0.018	0.0003 ± 0.0016
	1000	0.025 ± 0.019	0.0002 ± 0.0018
f_5	50	0.011 ± 0.011	$(2.0 \pm 223.6) \times 10^{-7}$
	250	0.011 ± 0.011	$(5.0 \pm 474.3) \times 10^{-7}$
	1000	0.011 ± 0.012	$(4.0 \pm 632.4) \times 10^{-7}$
f_6	50	0.0049 ± 0.0071	0.0 ± 0.0
	250	0.0049 ± 0.0071	0.0 ± 0.0
	1000	0.0049 ± 0.0078	0.0 ± 0.0
f_7	50	0.0025 ± 0.0047	0.0 ± 0.0
	250	0.0025 ± 0.0047	0.0 ± 0.0
	1000	0.0026 ± 0.0053	0.0 ± 0.0
f_8	50	0.445 ± 0.574	1.50 ± 0.22
	250	0.445 ± 0.572	1.49 ± 0.22
	1000	0.448 ± 0.570	1.48 ± 0.23
f_9	50	3.24 ± 2.38	-0.0012 ± 0.3978
	250	3.24 ± 2.37	0.002 ± 0.399
	1000	3.16 ± 2.24	-0.0029 ± 0.4044
f_{10}	50	15.8 ± 25.0	-0.63 ± 0.49
	250	15.7 ± 24.4	-0.626 ± 0.491
	1000	14.4 ± 19.7	-0.637 ± 0.504
f_{11}	50	1.33 ± 700.05	5.9 ± 913.9
	250	1.12 ± 141.10	0.066 ± 182.072
	1000	-0.44 ± 34.89	0.22 ± 46.05
f_{12}	50	0.62 ± 0.34	0.12 ± 0.09
	250	0.616 ± 0.337	0.124 ± 0.088
	1000	0.575 ± 0.318	0.125 ± 0.089
f_{13}	50	0.98 ± 0.04	0.97 ± 0.02
	250	0.914 ± 0.096	0.922 ± 0.039
	1000	0.731 ± 0.216	0.839 ± 0.077
f_{14}	50	0.017 ± 0.125	0.0049 ± 0.0031
	250	0.176 ± 0.308	0.027 ± 0.017
	1000	1.15 ± 0.68	0.114 ± 0.068

Table 4. Classification accuracy for observational *XMM* light curves of objects as listed in Table A1. Classifications are made on the basis of a cut at 50 per cent threshold probability for containing a QPE or not.

Time bin (s)	No. with QPE	No. without QPE	Accuracy	Purity	Completeness	F_1 score
50	12	57	0.319	0.203	1.0	0.338
250	12	54	0.212	0.188	1.0	0.316
1000	12	52	0.328	0.208	0.917	0.338

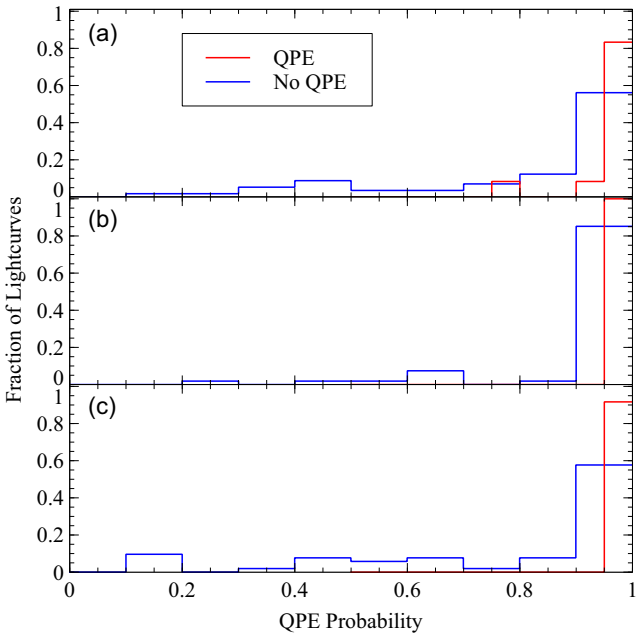

Figure 5. Probability that individual light curves in the observational data sample contain QPEs as per the results of the neural network classification. Panels show the results for light curves binned at (a) 50 s, (b) 250 s, and (c) 1000 s. The QPE and non-QPE samples displayed are normalized by the number of light curves in each sample.

several of the features the populations show differences, both for the simulated and observed samples. Features 1, 2, 3, 8, and 12 all show distinct differences between the QPE and non-QPE populations, with the distributions for the real observed light curves and the simulated light curves in each class clearly overlapping for all three time bin values. Fig. 9 shows how these five features interact in more detail. From this we can see that the combinations of f_1 and f_2, f_2 and f_3, f_2 and f_{12}, and f_8 and f_{12} appear to show a high level of separation between the two classes of simulated light curves.

Kolmogorov–Smirnov testing for all 14 features reveals that for the simulated feature distributions the classes are distinct at the 1 per cent confidence level for each individual feature at all-time bin values. In comparing the feature distributions for the two classes of real observations, Features 1, 2, 4, 8, 9, 10, and 13 were individually significantly different at the 1 per cent confidence level for all-time bin values. Features 3, 5, and 14 were all individually significantly different at a confidence level of 5 per cent for all three time bin values, and Feature 12 was significant at 10 per cent. Features 6 and 11 were also individually significantly different at the 10 per cent level for time binning of 50 s, but were not significantly different at the other two time bin values. Feature 7 did not show a significant difference at any value for time binning. Features 6 and 7 are, however, affected by the time binning of the simulated light curves, as all simulated light curves are of duration 100 ks, and so the number of points in the light curves decreases by a factor of twenty as the time binning increases from 50 to 1000 s. Additionally, these results with regards to the observational feature distributions should be taken with some caution due to the reduced sample size; the QPE class only contains 12 light curves binned at 1000 s.

If we consider networks trained for only individual features, and with architecture optimized as described in Section 2.4, we find that certain features at different values for the time bin appear to be more informative than others. With a time binning of 50 s, we achieve a classification accuracy of 0.870 and F_1 score of 0.609 using only Feature 3. With a time binning of 250 s, we achieve a classification accuracy of 0.970 and F_1 score of 0.917 using only Feature 8, or an accuracy of 0.894 and F_1 score of 0.588 using only Feature 12. With a time binning of 1 ks, we achieve a classification accuracy of 0.922 and F_1 score of 0.828 using only Feature 9.

4.2 Threshold probability optimization

Our default measure for distinguishing between light curves containing QPEs and not containing QPEs was to compare the classification probabilities, as the output of a SOFTMAX layer, and assign the class as that which has the highest probability. With our problem being that of a binary classification this means that initially we set a nominal threshold probability of 50 per cent between the two classes. Following the results of the classification of the observational data in Section 3.2 we then considered whether an adjustment to the level at which we split the two classes could improve the performance. We perform a simple grid search for threshold probability between 0 and 1 at steps of 0.000 001 with the results of the classification at all three neural networks and find that altering the level at which we delineate between light curves containing QPEs and those which do not can significantly improve the performance of the classifier. The full results of the optimization are given in Table 6, and the associated optimized confusion matrices are displayed in Fig. 10, 11 and 12. By varying the threshold probability at which we distinguish between QPE and non-QPE containing light curves we were able to make significant improvements in the accuracy and F_1 scores for all three data sets. In the case of all three classifiers, we can improve the accuracy by raising the required QPE probability. For each of the three classifiers, the optimal QPE probability was very close to 1.0, which is indicated by the distributions of light-curve probabilities as seen in Fig. 5. In all three cases by optimizing the accuracy we reduce the completeness of the classifier, but given that the sample being classified is biased towards light curves not containing QPEs the increase in false negative classifications is far outweighed by the decreases in each case of false positive classifications. For the 50 s classifier, 4 out of the 12 QPE containing light curves are misclassified at a higher probability, but an additional 46 of the 57 non-QPE light curves are then correctly classified, and we see a similar increase in the performance of the 250 s classifier. The greatest increase in performance is seen in the 1000 s classifier, where at the optimal threshold probability level for accuracy we reduce the misclassification rate for non-QPE containing light curves to 0 per cent.

After optimization with all three classifiers there is one false negative result which appears in all three samples, and one false positive which appears in both the 50 and 250 s binned data sets. Observation 0304 190 101 of NGC 1331 is classed as containing a QPE with probability 0.999 999, 0.999 976, and 0.999 329 across the 50, 250, and 1000 s classifiers, respectively. The light curve for observation 0304 190 101 of NGC 1331 is shown in Fig. 13. Observation 0861 910 301 of eRASSU J023147.2−102010 is classed as not containing a QPE with probability 0.986 134, 0.995 476, and 0.318 089 across the 50, 250, and 1000 s classifiers, respectively. The light curve for observation 0861 910 301 of eRASSU J023147.2−102010 is shown in Fig. 14.

If we look at the individual features, extracted from the 250 s binned light curves, which achieved the highest levels of classification accuracy, we find that in both cases many of the features for the erroneously classified light curves are more indicative of their incorrect classes. For observation 0304 190 101 of NGC 1331 Features 1, 9, and 10 were outliers for the non-QPE class, with

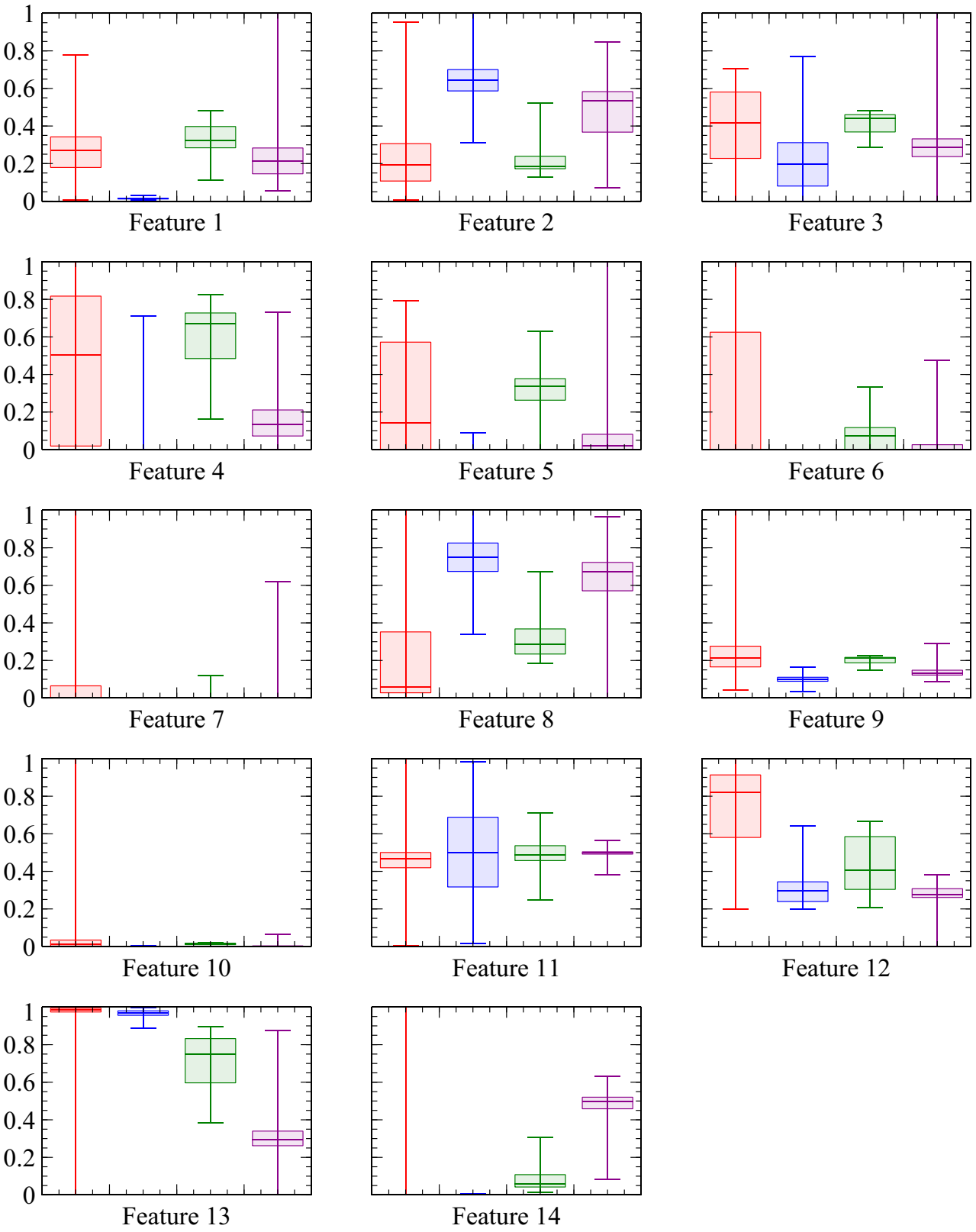

Figure 6. Distribution of feature values for the 14 features used in the classification derived from light curves binned at 50 s. Box plots display the distribution of features from left to right for the populations of: simulated light curves containing QPEs (red); simulated light curves without QPEs (blue); observed light curves containing QPEs (green); and observed light curves without QPEs (purple). Feature values have been normalized between 0 and 1.

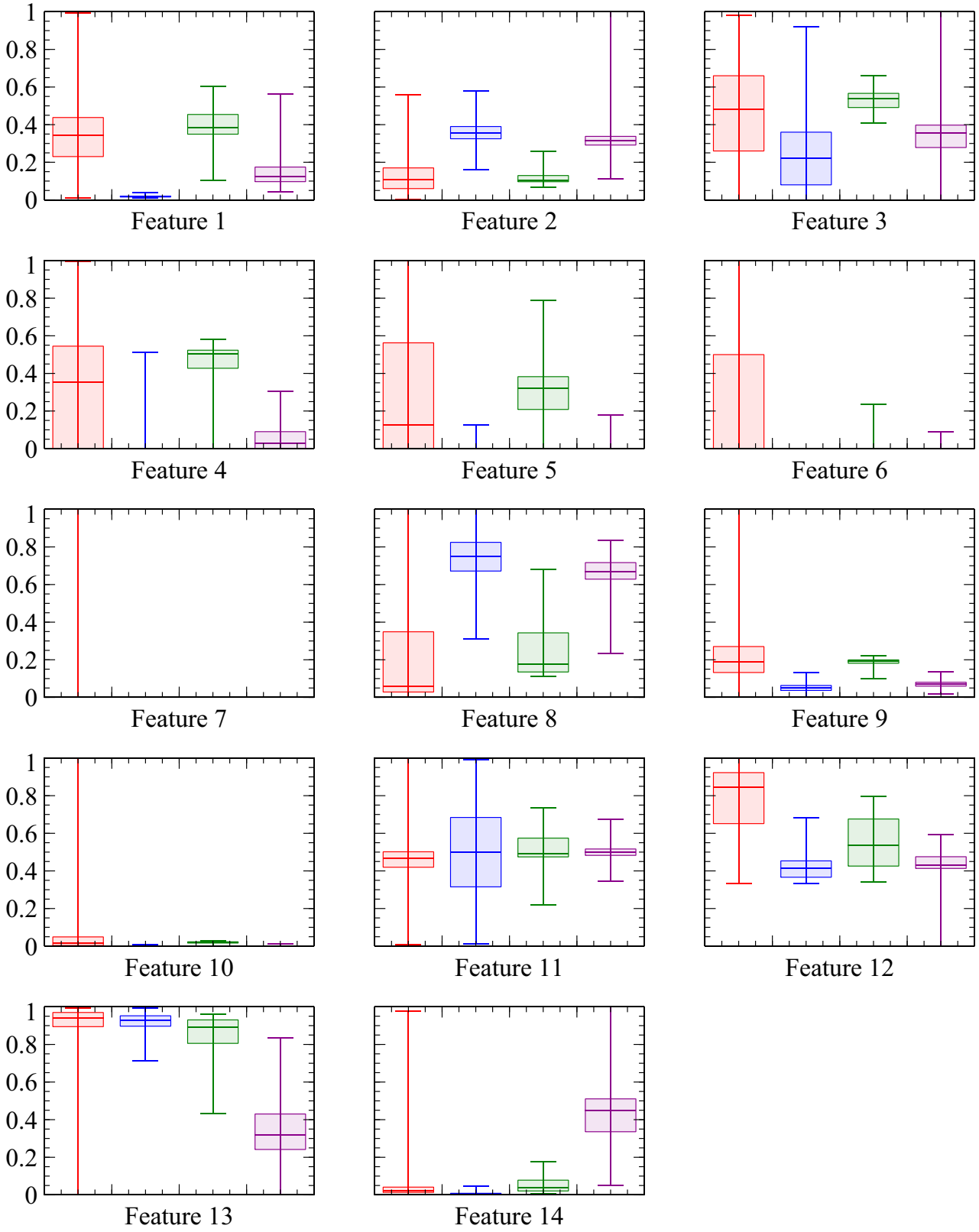

Figure 7. Distribution of feature values for the 14 features used in the classification derived from light curves binned at 250 s. Box plots display the distribution of features from left to right for the populations of: simulated light curves containing QPEs (red); simulated light curves without QPEs (blue); observed light curves containing QPEs (green); and observed light curves without QPEs (purple). Feature values have been normalized between 0 and 1.

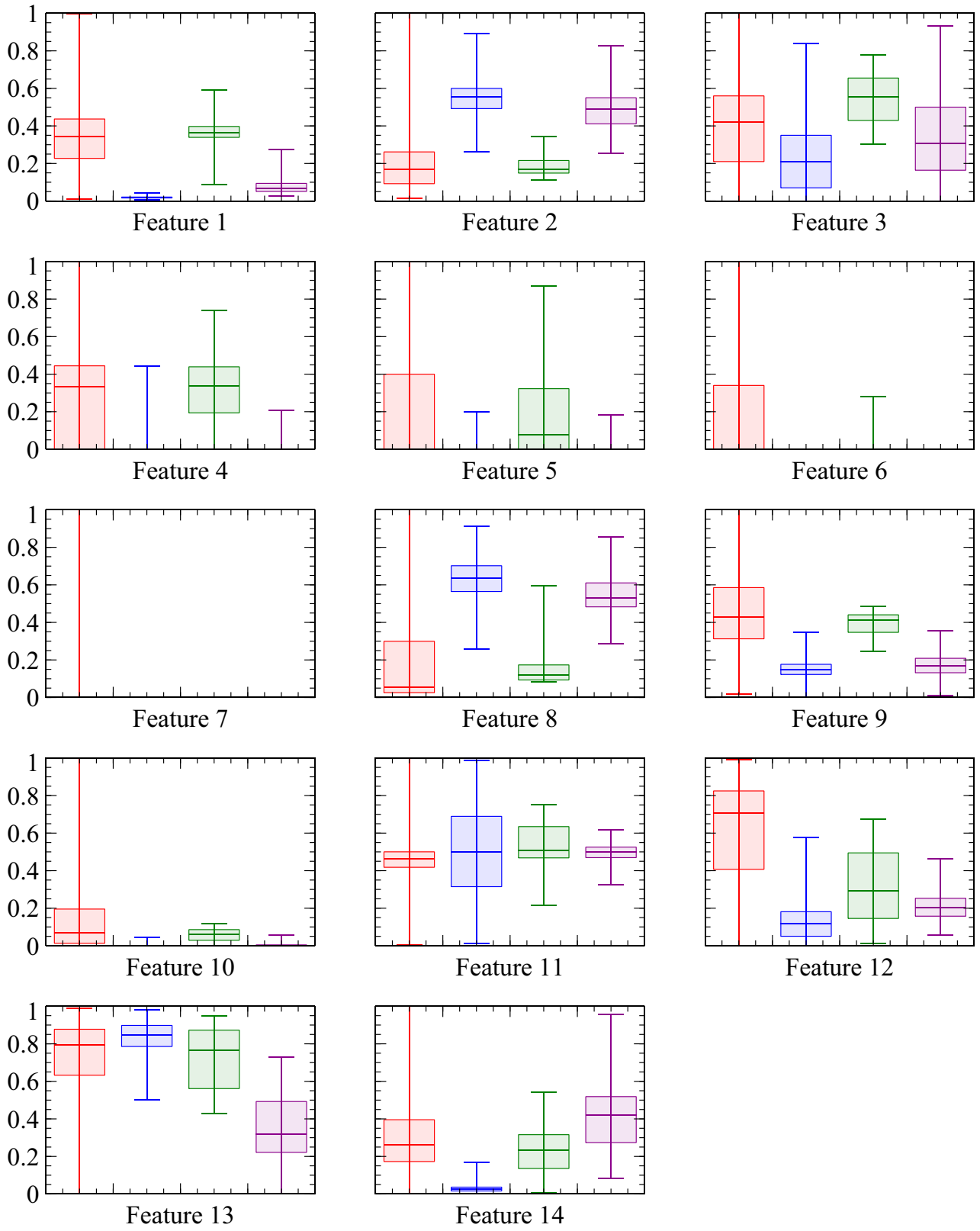

Figure 8. Distribution of feature values for the 14 features used in the classification derived from light curves binned at 1000 s. Box plots display the distribution of features from left to right for the populations of: simulated light curves containing QPEs (red); simulated light curves without QPEs (blue); observed light curves containing QPEs (green); and observed light curves without QPEs (purple). Feature values have been normalized between 0 and 1.

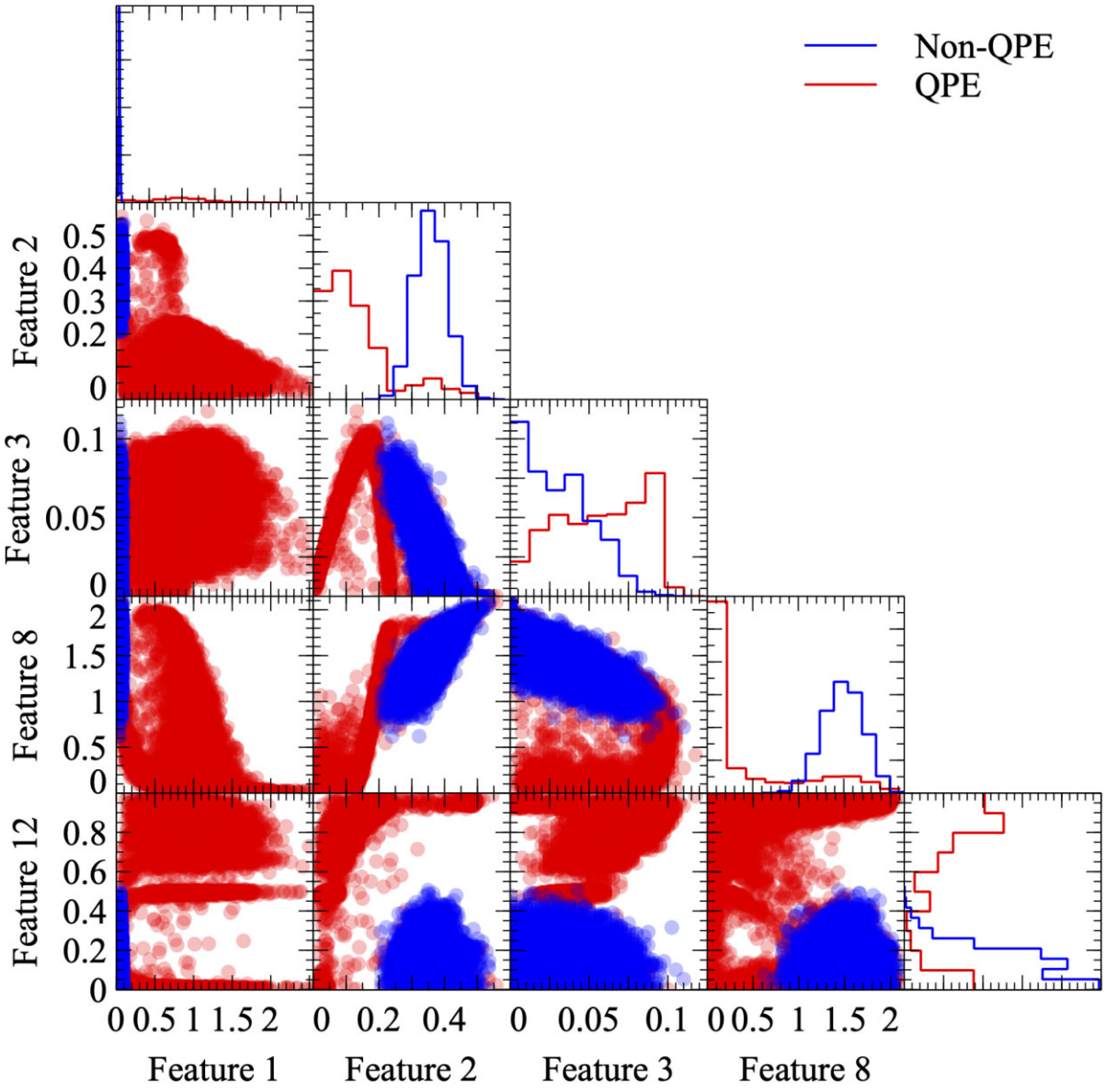

Figure 9. Distribution of feature values for Features 1, 2, 3, 8, and 12. For clarity the features for a sample of 10 000 of each of the simulated populations of light curves are shown in red (QPE) and blue (non-QPE). Histograms show the frequency density distribution of features for all 50 000 light curves in each class.

Features 1 and 9 showing significant differences between the two classes. For values of $f_1 = 1.52$ and $f_9 = 1.59$ this observation sits at 1.42σ (147.2σ) and 0.70σ (3.98σ) from the mean for the QPE (non-QPE) class. The average deviation from the class feature means for observation 0304 190 101 was 1.50σ for the QPE class, and 34.51σ for the non-QPE class. For observation 0861 910 301 of eRASSU J023147.2−102010 Features 4, 5, 6, 7, 8, 12, and 14 were outliers for the QPE class, with Features 4, 5, 6, 8, and 12 being the most extreme values for the QPE class. For values of $f_4 = 0.0, f_5 = 0.0, f_8 = 1.36$, and $f_{12} = 1.36$, this observation sits at 0.156σ (1.34σ), 0.011σ (0.935σ), 0.060σ (1.605σ), and 1.26σ (1.79σ) from the mean for the non-QPE (QPE) class. The average deviation from the class feature means for observation 0861 910 301 was 0.858σ for the QPE class, and 10.22σ for the non-QPE class. As the features overall for both of these observations would seem to indicate that both share more in common with the QPE class it is evident that the effects of individual features which may be more important than others are causing the misclassifications.

A particular issue in this classification is the high rate of false positive identifications, as QPEs are rare in appearance. If we look at those light curves which do not contain QPEs, we see a distinct difference in the average count rates for those which are correctly or incorrectly identified across all three time bin values. Those light curves which do not contain QPEs have average count rates of 0.402, 0.439, and 0.447 cts s^{-1} across the three time bin values when considering those which are available for classification. The average count rates for those observations which are categorized correctly (and incorrectly) are 0.472 (0.051) for the 50 s sample, 0.447 (0.005) for the 250 s sample, and 1.045 (0.103) cts s^{-1} for the 1000 s sample. Observation 0304 190 101, shown in Fig. 13, is misclassified with all three networks and has an average count rate of 0.0047 cts s^{-1}, indicating that even when binned at 250 s it is likely that there will be a high proportion of empty bins and as such the classification may well be affected by Poisson noise due to the low count rates.

Table 6. Optimal ranges for the cut in probability value at which we define a light curve as containing QPEs or not. Threshold probability ranges are closed intervals which give the optimized threshold probability values for accuracy, purity, completeness, and F_1 score.

Time bin (s)	Metric	Conf. range	Metric value
50	Accuracy	[0.999 206, 0.999 211]	0.8841
	Purity	[0.999 206, 0.999 211]	0.6429
	Completeness	[0, 0.778 044]	1.0
	F_1 score	[0.999 206, 0.999 211]	0.6923
250	Accuracy	[0.999 886, 0.999 944]	0.8939
	Purity	[0.999 977, 0.999 998]	1.0
	Completeness	[0.0, 0.986 412]	1.0
	F_1 score	[0.999 886, 0.999 893]	0.6666
1000	Accuracy	[0.999 330, 0.999 455]	0.9844
	Purity	[0.999 330, 1.0]	1.0
	Completeness	[0.0, 0.999 999]	1.0
	F_1 score	[0.999 330, 0.999 455]	0.9565

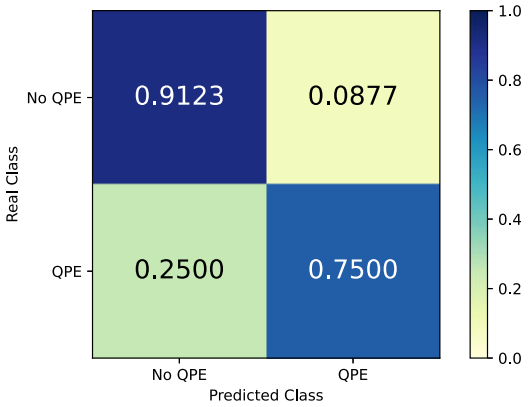

Figure 10. Confusion matrix showing the classification of 57 light curves without QPEs and 12 with light curves with QPEs present. The light curves were binned at a rate of 50 s, and their full details are given in A1. Threshold probability for requiring a QPE is set at the level which optimizes accuracy of 0.999 21.

4.3 Application to the XMM Serendipitous Source Catalogue

Having achieved high levels of accuracy in classifying simulated light curves, and with the moderate level of success achieved in classifying real observational data, we now consider whether any new QPE candidates can be identified through application of the networks to the XMM Serendipitous Source Catalogue.

We filter the XMM Serendipitous Source Catalogue (4XMM-DR12)[3] for observations which contain pn detections, have light-curve data which are created as described in Webb et al. (2020), and have total exposure times greater than 50 ks. Detections are referred to by their 'SRCID' tags as given in the 4XMM-DR12 data release. We also screen for any detections which could be spurious (possibly near an extended bright source, or where the detector coverage is particularly low) using the criteria that `SUM_FLAG=0`. From the total catalogue of 939 270 items this creates a sample of 83 531 detections for consideration. We download the pn light-curve data for these detections, and then screen these light curves for time bins of less than 50 s, and rebin the light curve to 50, 250, and 1000 s. Due to issues with very low-count-rate observations as identified in

[3] http://xmmssc.irap.omp.eu/Catalogue/4XMM-DR12/4XMM_DR12.html

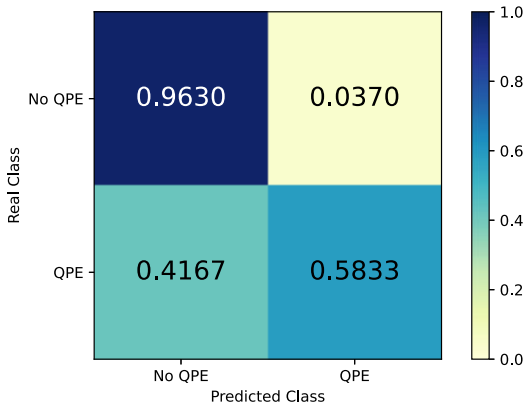

Figure 11. Confusion matrix showing the classification of 54 light curves without QPEs and 12 with light curves with QPEs present. The light curves were binned at a rate of 250 s, and their full details are given in A1. Threshold probability for requiring a QPE is set at the level which optimizes accuracy of 0.9999.

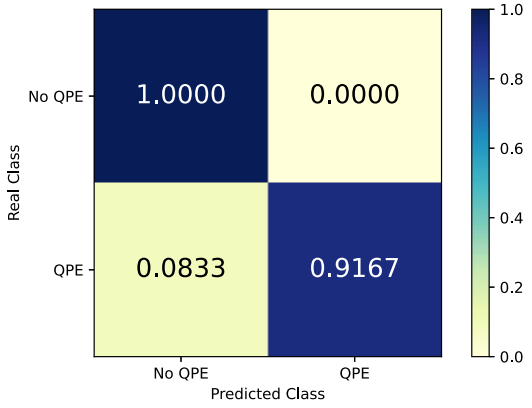

Figure 12. Confusion matrix showing the classification of 52 light curves without QPEs and 12 with light curves with QPEs present. The light curves were binned at a rate of 1000 s, and their full details are given in A1. Threshold probability for requiring a QPE is set at the level which optimizes accuracy of 0.9994.

Section 4.2, we then screen for any light curves with an average count rate below 0.01 cts s^{-1}. Finally, due to issues with flaring events which are common at the start or end of observations depending on their position during observing cycles we then remove the first and last 15 ks of each observation. Features are then extracted from each of the three light curves for each detection and the objects are classified according to the networks as trained in Section 2.4, and tested on the observational data as described in Section 3.2. As the phenomenon we are trying to find is very rare, we want to reduce the numbers of false positive events in any sample which is to be scheduled for follow-up by an expert observer. In Fig. 10, 11, and 12, we can see that with very high threshold probabilities we could still flag very large numbers of false positive results given the total sample size. As such, we select detections from the catalogue for further review based on the result of classification from all three networks. In order for a light curve to be flagged for further review, we require that it achieves a score of:

(i) 0.999 with features extracted from a 50 s binned light curve,
(ii) 0.9999 with features extracted from a 250 s binned light curve, and
(iii) 0.999 with features extracted from a 1000 s binned light curve.

Table 7. Predictions for the 11 observations of known QPE sources contained within the *XMM* Serendipitous Source Catalogue (4XMM-DR12). For each object we note the observation ID, the predictions from the three networks for the three time bin values (P50, P250, and P1000), and whether the light curve is flagged for further inspection as per the methodology in Section 4.3.

Object	OBSID	Eruptions	P50	P250	P1000	Flag
GSN 069	0823 680 101	2	0.999	0.9999	1.000	Y
–	0831 790 701	5	0.9986	0.9993	red0.999	N
–	0851 180 401	5	0.999	0.9999	0.999	Y
–	0864 330 101	4	0.829	0.9772	0.996	N
RX J1301.9+2747	0124 710 801	1	0.998 98	0.9995	1.000	N
–	0851 180 501	3	1.000	0.9995	1.000	Y
–	0864 560 101	8	0.998	0.9999	0.999	N
XMMSL1 J024916.6−041244	0411 980 401	1	0.768	0.9721	0.999	N
eRASSU J023147.2−102010	0861 910 201	2	0.998 98	0.9998	0.984	N
–	0861 910 301	1	0.952	0.9767	0.161	N
eRASSU J023448.9−441931	0872 390 101	9	1.000	0.9999	1.000	Y

At these threshold probabilities, we achieve purity (and completeness) scores of 0.6 (0.75), 0.86 (0.58), and 0.85 (0.92) with the 50, 250, and 1000 s classifiers, respectively, on our initial testing data, with 22 per cent, 11 per cent, and 20 per cent of the light curves in our sample being flagged at these threshold probability levels. With these selection criteria from the total sample of 83 531 detections, 705 light curves are identified for further examination with six being classified as containing QPEs with a probability of 1.0 across all three networks.

When this selection criteria is applied to the known observations of QPE candidates which are in the 4XMM-DR12 catalogue we find that for the 11 observations which are in the catalogue only four meet our selection criteria. Of those seven observations which do not meet the criteria, one only fails the criteria for one of the three networks. In all but one of the cases the predictions for the observations are strongly favoured towards containing QPEs, see Table 7 for the full results. The very low score for observation 0861 910 301 of eRASSU J023147.2−102010 when binned at 1 ks is an outlier. These results are not entirely unexpected due to the nature of the pre-processed light curves available with the XMM Serendipitous Source Catalogue. The light curves which are associated with the catalogue are for the full range of energies detected by *XMM–Newton* (0.2–12.0 keV), and we know from previous examination of QPE sources that the phenomenon is largely contained to lower energies (Miniutti et al. 2019; Giustini et al. 2020; Arcodia et al. 2021; Chakraborty et al. 2021, etc.). As such, the variability which is clear and evident in the 0.2–2.0 keV band appears suppressed and of a smaller amplitude when observed in the full 0.2–12.0 keV range, and it is not surprising that fewer of the known QPE sources meet our criteria for flagging in this case. Across the different time bin values, we observe changes in the features when considering the light curve for the full 0.2–12.0 keV energy band. Feature 1 for all observations decreased with a change to the energy band, Feature 9 displayed an average decrease across all three time bin values, and for 50 and 250 s binned curves Feature 8 displayed an average increase. All three of these changes are indicative of movement from QPE to non-QPE populations, and would explain the decreased probabilities of the classifications towards containing QPEs.

We then manually examined the light curves for the 705 detections which were flagged from the Serendipitous Source Catalogue. The first result of note is that a large proportion of the flagged detections came from observations which contained several detections flagged as containing QPEs. Of the 705 detections, 484 (69 per cent of all flagged detections) came from a combination of 37 observations, all of which contained at least five detections. Observation 0604 960 301 contained 46 detections which were flagged as containing QPEs at the conservative cuts implemented in the methodology described above. In the case of some of these observations, there are unusual patterns affecting the detections which are flagged as containing QPEs. One example is that of observation 0804 680 101, where there are five detections flagged as containing QPEs which all show a similar profile with peaks simultaneously in their light curves. See Fig. 15 for the 250 s binned light curves for sources 6, 7, 10, 11, and 22 from observation 0804 680 101. For this specific observation, there are peaks across the light curves of all five objects simultaneously at $\sim$15 ks, $\sim$24 ks, $\sim$75 ks, $\sim$100 ks, and $\sim$115 ks after the start of the observation. These features are evident in the light curves for all five sources, despite them being located on different parts of the detector, with source 6 being located on CCD 1, sources 7 and 10 being located on CCD 11, and sources 11 and 22 being located on CCD 10, and without overlapping raw positions on the detectors. The manner in which the light curves are created is standard across all detections in the catalogue. It requires the automatic detection of source locations, with background regions being drawn from empty regions on the same CCD as the source. Any other detections flagged in the background region are screened out. That sources located on different parts of the detector should have simultaneous, correlated variability is unlikely, and at present it is unknown as to what may be causing these features to appear across multiple detections in certain observations, but a manual recreation of the light curves for these detections appears to show that background features are not appropriately being subtracted.

The 705 detections were then grouped by observation ID, to allow for faster manual filtering by repeated features across detections in the same observation, and were then all reviewed manually. From this sample, we were able to identify 27 detections where their light curves warranted further examination. The details of these 27 detections are listed in Table B1. The detections were then cross-referenced with the SIMBAD astronomical data base (Wenger et al. 2000) to determine the names and types of sources to determine if the objects could be AGN as viable QPE sources. In the cases where there was no correlated source in SIMBAD, we then referenced the identities of those sources through NASA/IPAC Extragalactic Database. Of these sources, 19 were identifiable as stars or X-ray binary systems and not appropriate for further analysis to determine the presence of QPEs, and the remaining eight were galaxies or un-

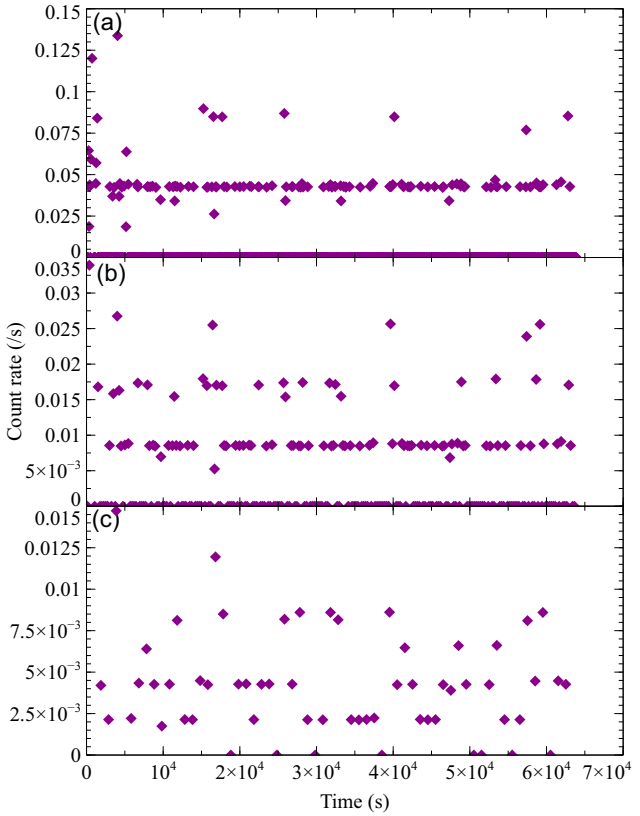

Figure 13. Light curve for observation 0304 190 101 of NGC 1331 which was misclassified as containing QPEs when binned at 50 and 250 s. The light curve is shown binned at rates of (a) 50 s, (b) 250 s, and (c) 1000 s and was misclassified in all cases. It appears likely that the misclassification is due to the effects of Poisson noise being more amplified given the low average count rate, which would be most noticeable for the two shorter time bin values.

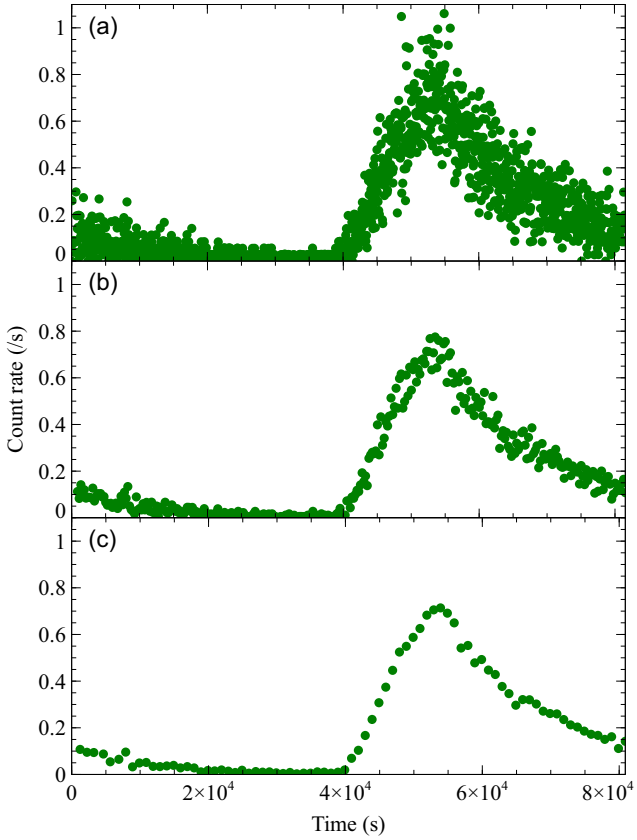

Figure 14. Light curve for observation 0861 910 301 of eRASSU J023147.2–102010 which was misclassified as not containing QPEs. The light curve is shown binned at rates of (a) 50 s, (b) 250 s, and (c) 1000 s and was misclassified in all cases.

classified sources. For those eight sources, we then downloaded and reprocessed the observation event lists, manually screened for flaring events, and created light curves for the 0.2–2.0 and 2.0–12.0 keV bands to determine whether the variability was present in soft X-ray bands. Of those eight sources, the variability was not present at lower energies for detections 200 823 401 010 018, 204 045 402 010 025, and 206 019 305 010 009. For the remaining five sources, we then created light curves in three narrower energy bands, 0.2–0.5, 0.5–1.0, and 1.0–2.0 keV, to determine whether the variability which appears in the soft X-rays follows the energy-dependent characteristics seen in confirmed QPE sources. At this stage, we ruled out detections 200 773 401 015 192, 201 068 601 010 017, 206 010 101 010 020, 207 277 801 010 019, and 208 235 803 010 011 as there is no significant difference seen in the variability in the three energy bands, with the flare being the same width, peaking at the same time and having the same amplitude against the quiescent rate in all three narrower bands. We give an example of a source which has been flagged and examined in further detail in Fig. 16. In this case, the source identified (SRCID 207 223 603 010 022) by the automatic classification pipeline and manual inspection was classified as a red dwarf, likely displaying chromospheric flares, and as such no further analysis is required. At this stage, our pipeline has not identified any new candidate QPE sources, with common sources of contamination being: instrumental background subtraction; chromospheric flares from red dwarfs; and flares from other classes of variable stars.

5 CONCLUSIONS

In this analysis, we set out to determine whether it would be possible to identify light curves containing QPEs by means of time-domain only features and commonly used machine learning classification techniques. This analysis used the freely available TENSORFLOW packages for PYTHON and features which had been previously used to classify other types of variability (Sokolovsky et al. 2017; Faisst et al. 2019), and were able to achieve accuracies of over 94 per cent in classifying simulated light curves and accuracies of over 98 per cent in classifying real observational data. The classification accuracy we have obtained is greater than that of Lo et al. (2014) (∼77 per cent) when only considering time-domain features and also uses a smaller number of features (14 rather than 27). The accuracy we have achieved is even greater than that obtained by Lo et al. (2014) (∼97 per cent) and Farrell et al. (2015) (∼92 per cent) when considering non-temporal features, which we do not consider in this analysis, in addition to time-domain features. We chose not to use spectral features as part of the network training process for two reasons: Simulating spectral features alongside temporal features would have been more computationally demanding; it could bias our classifier towards only finding QPE hosts with spectral features that are similar to those which have already been discovered. As all current QPE hosts have very soft spectra we could miss out on some harder X-ray sources of QPEs. As it is, the training on features sets takes of the order of minutes on a standard desktop computer, and the time taken for the creation of features and classification of light curves with the networks once trained is trivial.

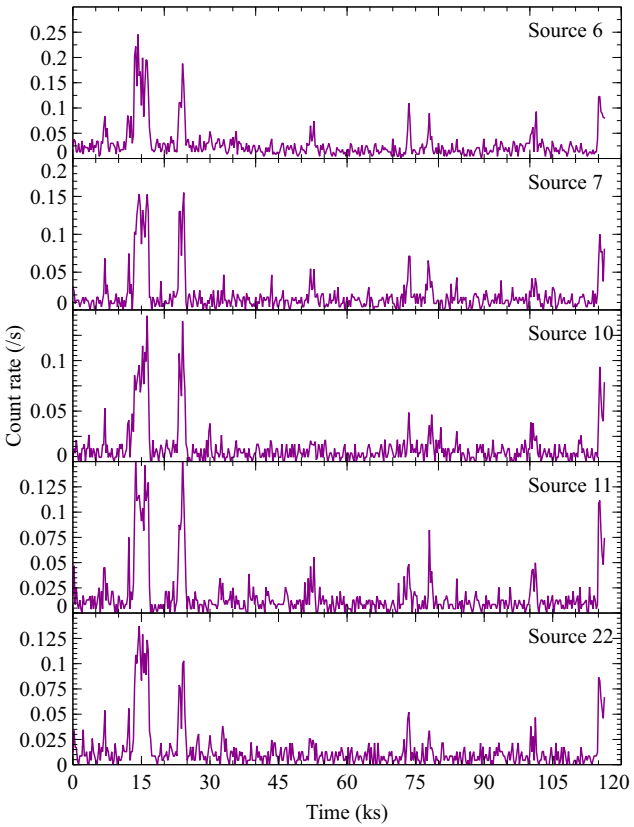

Figure 15. Light curves for five sources flagged as containing QPEs within the observation 0804 680 101. All light curves are binned at 250 s and show a consistent pattern of variability despite sources being located on different sections of the detector. Panels display (from top to bottom) the light curve for sources 6, 7, 10, 11, and 22.

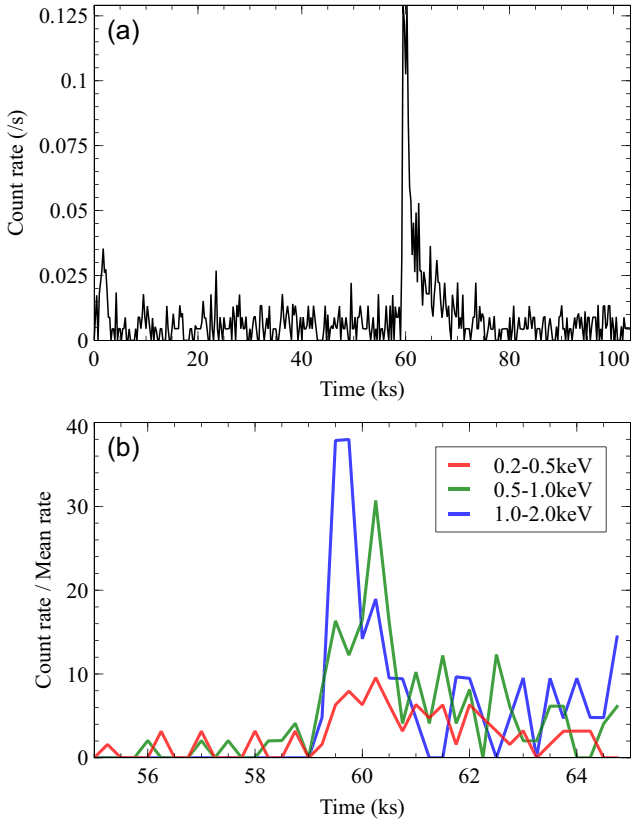

Figure 16. Light curves for the detection of 207 223 603 010 022 as source 22 in observation 0722 360 301. Top panel (a) shows the full band 0.2–12.0 keV light curve binned at a rate of 50 s for the source and the lower panel (b) shows the energy dependence of the flare located around 60 ks after the start of the observation binned at a rate of 250 s.

5.1 Further work

Further work could provide several ways to improve the efficacy of our classifier. Our approach to simulating light curves is itself limited by the number of eruptions which have currently been detected. With an increase in the number of known QPEs we could improve our understanding of the underlying distributions for the QPE features and so create a more accurate training data set. Our approach has also not required any information on the brightness of sources being considered, as this could be impacted by the sensitivity of different instruments. We should also consider the features being used in the classification process and whether the features created by the simulated data set adequately represent those seen in observational data. Feature 1 is likely to not be representative of real observational data for the reasons identified in Section 4.1. A further correction to the baseline of light curves after the QPE signals have been imposed could alter the effectiveness of and authenticity of this feature, but the issues caused by Poisson noise at very low count rates may mean that it would be better to eliminate the feature entirely. Features 13 and 14 also do not appear to follow the same distributions in the simulated and observational data sets, as shown in Figs 6, 7, and 8, and corrections to the generation process could be considered in order to make them more representative of the observational data. When we consider subsets of 13 of the 14 features, we find that improved performance in classifying the observational data sample is possible. With 1000 s binned light curves, we can achieve an accuracy of 96.9 per cent when Feature 3 is omitted. Further analysis of combinations of fewer features may identify the optimal approach moving forward. Additionally, we can consider data from instruments other than *XMM–Newton*. We have chosen not to include instrumental effects in generating light curves and features in order to allow this tool to be applied to data from other observatories, and including QPE features derived from observations of QPE sources with *Chandra* and *NICER* and other observatories may further improve the accuracy of these classifiers. An investigation into the effects of including Poisson noise in the populations of simulated light curves found a decrease in the ability of the classifiers to distinguish between light curves containing QPEs or not. The addition of noise increased the baseline variance of the light curves which made the two classes in training appear less distinct. However, at count rates typical of the data sets in which QPEs are detected, the degradation in performance was small. With the additional computational power and time which was required to include Poisson noise in simulating light curves we do not consider that it is likely to produce improvements in the classification accuracy in the future. Further applications of similar methods in other wavebands may allow for the detection of other types of astrophysical transient, and the network design with a SOFTMAX layer allows us to extend to such a multiclass problem. Additionally, it may be preferable when working with other catalogues of data to train the network using the F_β metric. In this instance, we found a slight decrease in the efficacy of the classifier when applied to real data when training with F_β, and the number of light curves flagged for manual inspection, when the networks were optimized for accuracy, in Section 4.3 was

manageable. With a larger catalogue it may be beneficial to train with F_β tuned to produce a smaller number of objects for manual inspection, even if the proportion of false positive results in that sample could be higher.

Ultimately, this analysis can serve as a benchmark for the detection of QPEs in X-ray light curves with machine learning. The classification has been achieved with high levels of accuracy in simulated light-curve sets and high, albeit lower, accuracy in real observational data. While we were not able to conclusively identify any new QPE sources in the XMM Serendipitous Source Catalogue we have identified two sources which show some QPE-like behaviour which could be followed up with further observation.

ACKNOWLEDGEMENTS

All figures in this work were created with VEUSZ and MATPLOTLIB.This work is supported by the UKRI Centre for Doctoral Training in Artificial Intelligence, Machine Learning & Advanced Computing, funded by grant EP/S023992/1. This research has made use of data obtained from the 4XMM *XMM–Newton* serendipitous source catalogue compiled by the *XMM–Newton* Survey Science Centre. This analysis used the specific version 4XMM-DR12. This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France. This research has made use of the NASA/IPAC Extragalactic Database (NED), which is funded by the National Aeronautics and Space Administration and operated by the California Institute of Technology. The authors would like to thank the reviewers for their constructive comments during the review process. In particular we would like to thank Reviewer #2 for their identification of two of the previously un-categorized sources flagged during the XMMSSC analysis.

DATA AVAILABILITY

All data used in this analysis are freely available in the *XMM–Newton* Science Archive (http://nxsa.esac.esa.int/nxsa-web/). All analysis tools and products are available at: https://github.com/robbie-webbe/ML_QPEs.

REFERENCES

Abadi M. et al., 2015, preprint (arxiv:1603.04467)
Arcodia R. et al., 2021, Nature, 592, 704
Carter J. A., Agol E., 2013, ApJ, 765, 132
Chakraborty J., Kara E., Masterson M., Giustini M., Miniutti G., Saxton R., 2021, ApJ, 921, L40
Chen X., Qiu Y., Li S., Liu F. K., 2022, ApJ, 930, 122
Faisst A. L., Prakash A., Capak P. L., Lee B., 2019, ApJ, 881, L9
Farrell S. A., Murphy T., Lo K. K., 2015, ApJ, 813, 28
Franchini A. et al., 2023, preprint (arXiv:2304.00775)(
Giustini M., Miniutti G., Saxton R. D., 2020, A&A, 636, L2
Gladney H. M., Dowden B. F., Swalen J. D., 1969, Anal. Chem., 41, 883
González-Martín O., Vaughan S., 2012, A&A, 544, A80
Grushka E., 1972, Anal. Chem., 44, 1733
Guo C., Pleiss G., Sun Y., Weinberger K. Q., 2017, preprint (arXiv:1706.04599)
Hübner M., Huppenkothen D., Lasky P. D., Inglis A. R., 2022, ApJS, 259, 32
Ingram A., Motta S. E., Aigrain S., Karastergiou A., 2021, MNRAS, 503, 1703
Kaur K., Stone N. C., Gilbaum S., 2022, preprint (arXiv:2211.00704)
King A., 2020, MNRAS, 493, L120
King A., 2022, MNRAS, 515, 4344
Krolik J. H., Linial I., 2022, ApJ, 941, 24
Linial I., Metzger B. D., 2023, preprint (arXiv:2303.16231)
Linial I., Sari R., 2023, ApJ, 945, 86
Lintott C. J. et al., 2008, MNRAS, 389, 1179
Lo K. K., Farrell S., Murphy T., Gaensler B. M., 2014, ApJ, 786, 20
Lu W., Quataert E., 2022, preprint (arXiv:2210.08023)
Metzger B. D., Stone N. C., Gilbaum S., 2022, ApJ, 926, 101
Miniutti G. et al., 2019, Nature, 573, 381
Miniutti G., Giustini M., Arcodia R., Saxton R. D., Read A. M., Bianchi S., Alexander K. D., 2023, A&A, 670, A93
Musoke G., Liska M., Porth O., van der Klis M., Ingram A., 2023, MNRAS, 518, 1656
Muthukrishna D., Narayan G., Mandel K. S., Biswas R., Hložek R., 2019a, PASP, 131, 118002
Muthukrishna D., Parkinson D., Tucker B. E., 2019b, ApJ, 885, 85
Muthukrishna D., Mandel K. S., Lochner M., Webb S., Narayan G., 2022, MNRAS, 517, 393
Narayan G. et al., 2018, ApJS, 236, 9
O'Malley T., Bursztein E., Long J., Chollet F., Jin H., Invernizzi L., 2019, https://github.com/keras-team/keras-tuner(accessed 2022 April 2)
Pan X., Li S.-L., Cao X., Miniutti G., Gu M., 2022, ApJ, 928, L18
Panagiotou C., Kara E., Dovciak M., 2022, ApJ, 941, 57
Pashchenko I. N., Sokolovsky K. V., Gavras P., 2018, MNRAS, 475, 2326
Pedregosa F. et al., 2011, J. Mach. Learn. Res., 12, 2825
Raj A., Nixon C. J., 2021, ApJ, 909, 82
She R., Ho L. C., Feng H., 2017, ApJ, 835, 223
Shin M.-S., Sekora M., Byun Y.-I., 2009, MNRAS, 400, 1897
Sniegowska M., Czerny B., Bon E., Bon N., 2020, A&A, 641, A167
Sokolovsky K. V. et al., 2017, MNRAS, 464, 274
Suková P., Zajacek M., Witzany V., Karas V., 2021, ApJ, 917, 43
Timmer J., Koenig M., Timmer J., Koenig M., 1995, A&A, 300, 707
Wang M., Yin J., Ma Y., Wu Q., 2022, ApJ, 933, 225
Webb N. A. et al., 2020, A&A, 641, A136
Webbe R., Young A. J., 2023, MNRAS, 518, 3428
Wenger M. et al., 2000, A&AS, 143, 9
Wozniak P. R., 2000, Acta Astron., 50, 421
Xian J., Zhang F., Dou L., He J., Shu X., 2021, ApJ, 921, L32
Zhao Z. Y., Wang Y. Y., Zou Y. C., Wang F. Y., Dai Z. G., 2022, A&A, 661, A55

APPENDIX A: OBSERVATIONS USED FOR TESTING

Table A1 contains a list of all *XMM* observations of AGN which were processed and used in the real data testing phase. In all cases, the pn light curves were binned at 50, 250, and 1000 s to account for shot-noise in lower count-rate observations. Mass estimates are either taken from or the CHANSNGCAT catalogue.

Table A1. Details of observations of low-mass AGNs which were identified as part of a manual search for QPE candidates. Values for M_{BH} are given as $\log(M_{BH}/M_{\odot})$ and are as per CHANSNGCAT. Those observations which could not be rebinned at 250 s are denoted with (∗), and those which could not be rebinned at 1000 s are denoted with (†) against their observation ID numbers.

AGN name	M_{BH}	OBSID	Exp. (ks)
2XMM J123103.2+110648	4.87	0145 800 101	107.0
–	–	0306 630 101	80.9
–	–	0306 630 201	99.5
eRASSU J023147.2−102010	5.78	0861 910 201	94.2
–	–	0861 910 301	90.2
eRASSU J023448.9−441931	4.96	0872 390 101	95.0
–	–	0893 810 501	25.0
GSN 069	5.99	0657 820 101	14.9
–	–	0740 960 101	95.1
–	–	0823 680 101	63.3
–	–	0831 790 701	141.4
–	–	0851 180 401	135.4
–	–	0864 330 101	141.0
–	–	0864 330 201	133.1
–	–	0864 330 301	133.2
–	–	0864 330 401	136.1
NGC1331	5.48	0304 190 101	65.9
NGC3367	5.64	0551 450 101	31.4
NGC3599	5.85	0411 980 101	6.9
–	–	0556 090 101	43.6
NGC4467	5.86	0112 550 601	24.6
–	–	0510 011 501 (†)	10.1
–	–	0761 630 101	118.0
–	–	0761 630 201	118.0
–	–	0761 630 301	117.0
NGC4476	5.58	0200 920 101	109.3
–	–	0551 870 401	21.6
–	–	0551 870 601	21.3
–	–	0603 260 201	17.9
–	–	0803 670 501	132.0
–	–	0803 670 601	65.0
–	–	0803 671 001	63.0
–	–	0803 671 101	131.9
NGC4559	5.14	0152 170 501	42.2
–	–	0842 340 201	75.4
NGC4654	5.07	0651 790 201	28.9
NGC5273	5.97	0112 551 701	17.1
–	–	0805 080 401	110.9
–	–	0805 080 501	28.0
NGC6946	5.43	0093 641 501 (∗)	8.6
–	–	0093 641 601 (∗)	10.1
–	–	0093 641 701 (∗)	11.3
–	–	0200 670 101	16.4
–	–	0200 670 201 (†)	14.4
–	–	0200 670 301	15.6
–	–	0200 670 401	21.2
–	–	0401 360 101	20.9
–	–	0401 360 201	24.4
–	–	0401 360 301	24.4
–	–	0500 730 101	31.9
–	–	0500 730 201	37.3
–	–	0691 570 101	119.3
–	–	0794 581 201	50.0
–	–	0870 830 101	17.9
–	–	0870 830 201	17.7
–	–	0870 830 301	16.0
–	–	0870 830 401	17.8
NGC7314	5.59	0111 790 101	44.7
–	–	0311 190 101	83.9
–	–	0725 200 101	140.5
–	–	0725 200 301	132.1
–	–	0790 650 101	65.0
NGC925	6.0	0784 510 301	50.0
–	–	0862 760 201	42.0
RX J1301.9+2747	6.65	0124 710 801	29.8
–	–	0851 180 501	48.4
–	–	0864 560 101	134.9
XMMSL1 J024916.6−041244	5.29	0411 980 401	11.7
–	–	0891 800 601	33.8

APPENDIX B: TOP CANDIDATES FROM XMM SSC ANALYSIS

Table B1 containsp details of those detections from the XMM Serendipitous Source Catalogue (4XMM-DR12) which were manually identified as containing significant variability and where follow-up analysis was performed. Table B1 lists the type of object corresponding to the detection and whether it contains QPEs if a galaxy.

Table B1. Details of detections from the XMM SSC which were flagged as containing QPEs and were followed up with manual inspection. The last three columns contain details as to the type of astrophysical object, whether the variability is seen when only the 0.2–2.0 keV band light curve is created, and whether the energy dependent and temporal characteristics are indicative of QPE behaviour. Sources are listed in order of their 4XMM Catalogue source IDs.

SRCID	OBSID	Source No.	Object name	Type	Soft variability	QPE-like?
200 287 402 010 079	0655 050 101	11	2XMM J001527.9−390507	Star	N/A	N/A
200 560 203 010 005	0801 610 101	4	[BHR2005] 832-14	Low-mass star	N/A	N/A
200 669 401 010 002	0501 790 101	2	Cl∗ NGC 2547 JND 13-98	Star	N/A	N/A
200 669 401 010 075	0501 790 101	15	Cl∗ NGC 2547 NTJ 7-2323	Low-mass star	N/A	N/A
200 773 401 015 192	0555 630 401	136	2XMM J150122.1−414227	X-ray source	Y	N
200 823 401 010 018	0203 560 201	12	SDSS J111740.11+074411.7	Quasar	N	N/A
201 068 601 010 017	0803 990 301	4	[CPS95] X-12	X-ray source	Y	N
201 080 604 010 030	0555 780 601	5	[LBX2017] 780	High proper motion star	N/A	N/A
–	0604 960 801	3	–	–	–	–
–	0604 961 201	5	–	–	–	–
201 111 202 010 026	0305 540 701	14	[GY92] 463	T-Tauri star	N/A	N/A
201 111 202 010 048	0800 031 001	3	[GY92] 259	Young stellar object	N/A	N/A
201 125 903 010 151	0403 200 101	85	V∗ V1320 Ori	BY Dra variable	N/A	N/A
201 428 001 010 001	0142 800 101	1	X LMC X-4	High-mass XRB	N/A	N/A
202 010 902 010 004	0821 240 301	3	WISEA J020621.12−002346.7	Star	N/A	N/A
204 045 402 010 025	0674 050 101	34	2CXO J202103.2+365423	X-ray source	N	N/A
206 006 901 010 002	0600 690 101	2	UCAC4 509-131194	High proper motion star	N/A	N/A
206 010 101 010 020	0844 860 101	8	XMMU J004705.9−205239	X-ray source	Y	N
206 019 305 010 009	0601 930 501	6	WISEA J213740.24+002048.0	IR source	N	N/A
206 048 602 010 005	0604 860 201	5	WISEA J181323.39−325230.9	Red dwarf	N/A	N/A
206 939 901 010 022	0693 990 301	10	WISEA J111949.13+065305.6	Star	N/A	N/A
207 216 201 010 018	0721 620 101	18	2MASS J08384128+1959471	Eruptive variable star	N/A	N/A
207 223 603 010 022	0722 360 301	22	WISEA J220310.58−344406.7	Red dwarf	N/A	N/A
207 277 801 010 019	0790 380 901	18	[ELK2021] 22	Galaxy	Y	N
207 437 002 010 001	0743 700 201	1	TYC 4682-1697-1	Star	N/A	N/A
207 810 401 010 007	0781 040 101	7	SDSS J032048.68+003234.0	Low-mass star	N/A	N/A
208 235 803 010 011	0823 580 301	11	WISEA J124901.51−410131.6	IR source	Y	N

This paper has been typeset from a TEX/LATEX file prepared by the author.