

Background subtraction and transient timing with Bayesian Blocks

H. Worpel and A. D. Schwope

Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, 14482 Potsdam, Germany
e-mail: hworpel@aip.de

Received 23 February 2015 / Accepted 16 March 2015

ABSTRACT

Aims. We aim to incorporate background subtraction into the Bayesian Blocks algorithm so that transient events can be timed accurately and precisely even in the presence of a substantial, rapidly variable background.

Methods. We developed several modifications to the algorithm and tested them on a simulated *XMM-Newton* observation of a bursting and eclipsing object.

Results. We found that bursts can be found to good precision for almost all background-subtraction methods, but eclipse ingresses and egresses present problems for most methods. We found one method that recovered these events with precision comparable to the interval between individual photons, in which both source- and background-region photons are combined into a single list and weighted according to the exposure area. We also found that adjusting the Bayesian Blocks change points nearer to blocks with higher count rate removes a systematic bias towards blocks of low count rate.

Key words. methods: numerical – binaries: eclipsing – X-rays: binaries

1. Introduction

The detection and precise timing of transient events is of importance in all areas of astrophysics. In some applications, for instance X-ray astronomy, the desired timing accuracy is comparable to the interval between the arrival times of individual photons. In this situation it is necessary to identify unambiguously which photons mark the end of the pre-transient phase and the beginning of the post-transient phase and, if possible, estimate without systematic bias where between these photons the change occurs.

One method that has attracted recent interest is the Bayesian Blocks algorithm (Scargle 1998; Scargle et al. 2013). This algorithm may take as its input a list of photon arrival times, such as those produced by X-ray telescopes. The algorithm can also handle other data modes, but we here exclusively consider time tagged event data. The observation is therefore divided into cells, which are intervals containing a single photon, and periods of time of a nominally constant count rate are produced by finding an optimum number and placement of blocks of consecutive cells. Given a prior on the number of blocks, the algorithm finds an objectively best-fit set of blocks describing the observation, and within each block, the arrival times of the photons are consistent with a Poisson process with constant rate. The Bayesian Blocks algorithm is ideal for finding and accurately timing transients, because a change point between blocks, typically placed half-way between the last event of one block and the first event of the next, will be found, if and only if, the rate of arrival of photons changes detectably. Transient timing can therefore, in principle, be performed with a precision comparable to the interval between individual photons, and clearly no better precision is possible. This property makes the Bayesian Blocks algorithm an attractive candidate for transient timing.

Other binning methods exist that, similarly to the Bayesian Blocks method, place bin edges at locations only where the data justify it (e.g., Knuth 2006; Bélanger 2013), but we did not

consider them here. See also Burgess (2014) for a comparison of binning methods, in the context of gamma ray burst timing.

The motivation for this paper is to achieve accurate and precise eclipse timings of cataclysmic variables (CVs) and low-mass X-ray binaries (LMXB) using data from the *XMM-Newton* X-ray observatory. Much of the paper is presented with these applications in mind, but the methods and conclusions developed are not specific to this astronomical field and are not even limited to astronomy. We consider two effects that may affect accurate timing-measurements using the Bayesian Blocks algorithm.

Firstly, is the location of the cell edges between photons optimal? While most applications place the cell edges half way between two photons (e.g., Scargle 1998; Scargle et al. 2013; Ivezić et al. 2014), there is no reason why the edge cannot be placed at any point between the photons. In Sect. 2 we describe a systematic bias affecting the half-way placement and suggest an adjustment that removes this bias almost entirely. We also consider the case of a continuously varying source intensity, such as those caused by an extended emitting spot moving into or out of eclipse.

Secondly, *XMM-Newton* observations are frequently affected by soft proton flares originating in the terrestrial magnetosphere that contribute a substantial, and often rapidly variable, background (Lumb et al. 2002) that needs to be removed to recover the true variability of the source. Our data therefore consist of photon lists: one extracted from a region of sky surrounding the source and containing both source and background photons, and one taken from a source-free region of sky containing background alone. It is not obvious how to perform background subtraction using the Bayesian Blocks algorithm. Removing individual source region photons according to the background rate has been suggested (e.g., Stelzer et al. 2007), but we would prefer not to discard any data. An attempt at weighting individual photons in a manner very similar to what we develop here has recently been applied successfully by Mossoux et al. (2015). In Sect. 3 we generate theoretical source and background light

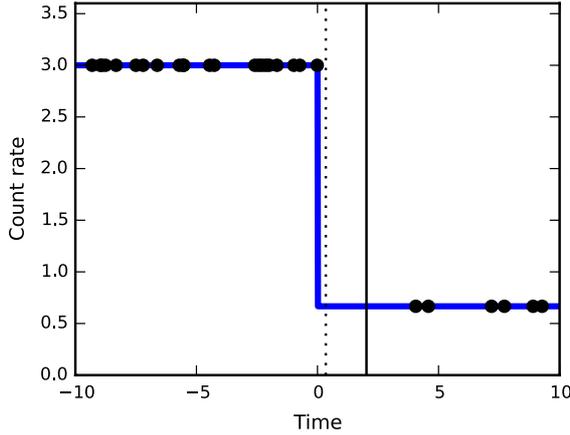


Fig. 1. Illustration of the change point bias toward blocks with lower count rate. The two count rates, 3 and 2/3, are shown as the blue line, with events placed according to a Poisson process shown as circles. The half-way point between the two terminal photons is shown as a solid vertical line, while the change point placed according to Eq. (3) is shown as a vertical dotted line. It is clear that the half-way change point method is biased towards blocks with lower count rate, but that the adjusted change point location gives a better estimate of the true location of the change point.

curves that contain transient events occurring at known times. In Sect. 4 we propose several modifications to the Bayesian Blocks algorithm to allow for background subtraction and evaluate them for their ability to recover the transients in the model light curve in the presence of a strong, rapidly varying, background.

2. Error analysis

We here used the geometric prior suggested by Scargle et al. (2013). This prior contains one adjustable parameter p_0 , that reflects the balance between suppressing spurious change points and retaining genuine ones. Throughout this paper we have taken $p_0 = 0.01$, which selects against false positives at the 99% level.

Most implementations of the Bayesian Blocks algorithm place the cell boundaries at the midpoint between two successive events (e.g., Scargle 1998; Scargle et al. 2013; Ivezić et al. 2014). This will bias the location of change points slightly towards blocks of low count rate. Consider a Poisson process that has rate r_0 for times $t < 0$ and rate r_1 for $t > 0$, with $r_1 < r_0$. We assume the last event of the first segment and the first event of the second segment occur at times t_0 and t_1 . Since they are both Poisson processes, t_0 and t_1 have mean values of $-1/2r_0$ and $1/2r_1$, and the Bayesian Blocks change point, since it falls halfway between the two, has a mean value $(r_0 - r_1)/(2r_0r_1)$. This is positive, so the location of the change point is biased toward the block with the lower count rate. Figure 1 illustrates this bias. For applications such as eclipse timings in the study of cataclysmic variables, this bias can be a severe drawback, because the count rate during the eclipse is probably very low. This bias will cause the eclipse to appear shorter than it really is. Furthermore, if the pre- and post-eclipse blocks are different in count rate, the eclipse midpoint may be shifted.

It is possible to improve the positioning of the change points as follows: we assume that the Bayesian Blocks algorithm has found a block of high count rate beginning at t_s , that contains n_0 events, and concludes with an event recorded at time t_0 , followed by a block of lower count rate. We wish to find the best location of the change point t_{cp} . Since we assume a Poisson process, t_0 is

exponentially distributed with standard deviation $1/r_0$ where r_0 is the count rate in the block. It follows that on average,

$$t_{cp} = t_0 + \frac{1}{2r_0}. \quad (1)$$

We also have

$$r_0 = \frac{n_0}{t_{cp} - t_s}, \quad (2)$$

and these can be solved simultaneously to give

$$t_{cp} = \frac{2n_0t_0 - t_s}{2n_0 - 1}. \quad (3)$$

The case of a rising count rate is very similar.

We tested this new adjustment with a simulated event series consisting of 100 events with count rate 3, up to $t = 0$, followed by another 100 events with count rate 2/3, and found the change point according to both the half-way method and the one given by Eq. (3). We repeated this test for 50 000 realizations of the series. The count rates here are dimensionless because the unmodified Bayesian Blocks algorithm is invariant under changes in the unit of time. The mean location of the recovered change point using the half-way change point location was $t_{cp} = 1.36 \pm 1.83$, comparable to the average spacing of photons in the lower count rate segment and indicating a substantial bias toward the lower count rate. For the new method, the mean location was $t_{cp} = 0.461 \pm 1.58$, an improvement by a factor of three. The small bias that remains is entirely due to the algorithm occasionally mistaking the first event of the second segment for an event in the first segment, if it happens to be placed close to the true change in count rate. The converse, mistaking an event in the higher count rate segment for one in the lower count rate segment, is also possible but much less likely since this requires several photons in the higher count rate segment to be poorly placed.

We repeated the experiment, discarding all trials with misidentified photons, and found that the mean change point locations for the half-way and adjusted change point locations were 0.66 ± 0.45 and -0.044 ± 0.20 . The adjusted change point method essentially has no bias, except for the uncertainty in determining which photons belong to which blocks.

Adjusting the location of change points by this method is only useful if the uncertainty in their location is smaller than the reduction in bias. To investigate under which circumstances the adjustment is useful, we performed the following tests. We produced many segmented light curves as above, consisting of 100 events with count rate $r_1 = 1$, followed by 100 events of count rate r_0 between 0.001 and 1. The division between the two count rates was at $t = 0$. We produced 1000 realizations of each of these light curves and ran the Bayesian Blocks algorithm on them, with and without the new change point adjustment, repeating trials where no change point was found at all. Then we compared the improvement in bias achieved (i.e., the difference of their absolute magnitudes) and compared it to the remaining uncertainty. The results are shown in Fig. 2.

The change point adjustment provides an improvement in bias greater than the remaining uncertainty for r_0 smaller than about 0.15, but no useful improvement for higher r_0 . This behaviour can be understood by considering the sources of uncertainty and bias for the two methods. For the unadjusted method, both bias and uncertainty are dominated by the placement of individual photons in the lower count rate segment. A misidentified photon contributes relatively little to the bias because, to

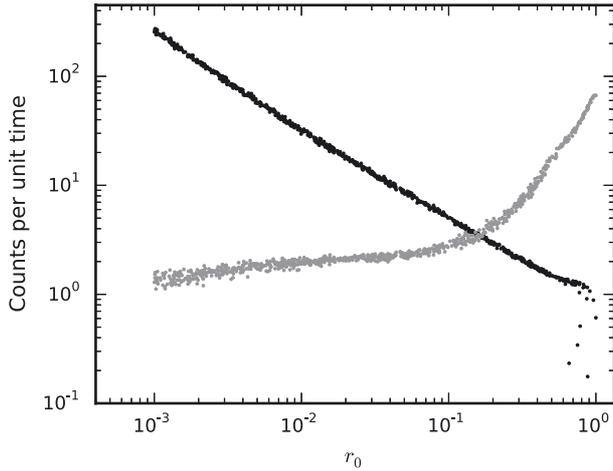


Fig. 2. Reduction in bias by employing the change point adjustment (black points), and remaining uncertainty in change point location (grey points), plotted against the lower of the two count rates. The improvement is larger than the remaining uncertainty for $r_0 < 0.15$.

be misidentified, it must be placed within about $1/r_1$ of the change in count rate, and this is smaller than $1/r_0$. For the adjusted method the bias consists almost entirely of misidentified photons, and the uncertainty is partly due to misidentified photons and partly due to the placement of individual photons in the r_1 segment. It follows that the bias and uncertainty in the unadjusted method decreases with increasing r_0 , because $1/r_0$ decreases, and both the bias and uncertainty in the adjusted method increase because misidentified photons become increasingly likely.

The adjustment of change points according to Eq. (3) is therefore most useful when one of the blocks is less than 15% as intense as the other. For the remainder of the paper, we use this method to determine the location of the change points.

2.1. Continuously varying count rate

In studies of eclipsing CVs, it is useful to determine or constrain the size of the emitting accretion spot on the white dwarf. When the spot passes behind the secondary star, its observed flux steadily decreases to zero. If this decrease is gradual enough, the Bayesian Blocks algorithm will produce blocks of intermediate count rate during the ingress or egress. If the decrease is too rapid, the algorithm will not resolve it and only produce an instantaneous change in count rate; the failure can still be used to constrain the size of the emitting region.

We investigated this issue by simulating many hypothetical eclipse egresses, with a very low eclipse count rate of 10^{-3} and post-eclipse count rates R of between 1 and 15. The egress itself was modelled as a linearly rising count rate with duration Δt between 0 and 50, and the fiducial time $t = 0$ was placed at the midpoint of the egress. For each set of R and Δt we made 1000 realizations and found the Bayesian Blocks change points. Figure 3 shows the probability that the egress will be resolved, that is, that the Bayes Blocks algorithm finds more than one change point in the egress. It is clear that this probability increases with increasing egress length and with increasing post-eclipse count rate, as one would expect.

The usefulness of this analysis for eclipse timing is clear. If the ingress or egress is not resolved, the size of the emitting spot can be constrained by finding the smallest Δt for which that ingress or egress would have been resolved. For instance, if an

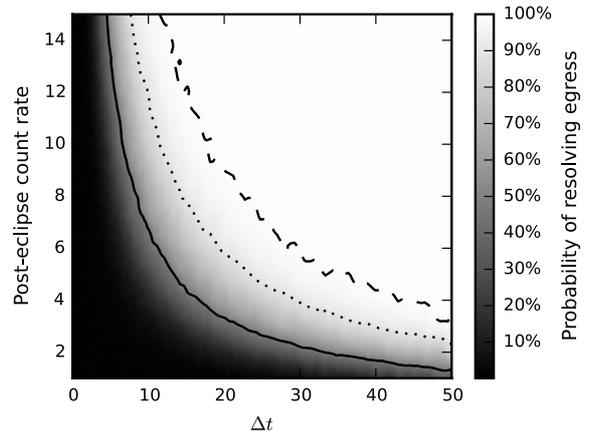


Fig. 3. Probability of resolving an eclipse egress with post-eclipse count rate R and eclipse egress duration Δt . Contours are shown at the 50%, 90%, and 99% significance level (solid, dotted, and dashed lines respectively). The probability of finding more than one change point during egress increases with increasing duration and post-eclipse count rate.

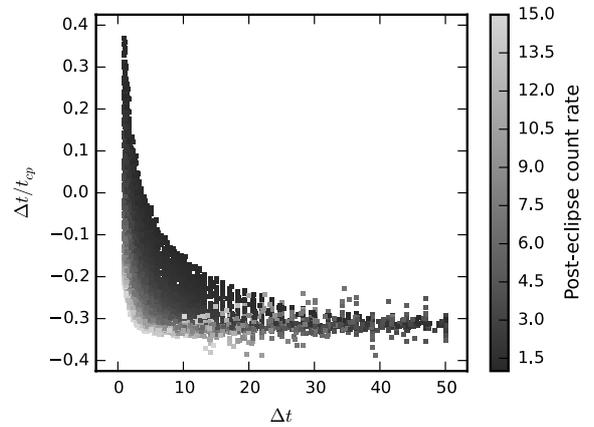


Fig. 4. Bias in the eclipse egress change point location as a fraction of egress duration, for simulations in which the egress was not resolved. Colours indicate the post-eclipse count rate.

eclipse egress in a real observation is not resolved and the post-eclipse count rate is 6 photons/s, it can be seen from Fig. 3 that the egress has a shorter duration than 30 s with approximately 99% confidence.

For simulations where the egress was not resolved, that is, where only one change point was found for the egress, we found the mean position of that change point. The results are shown in Fig. 4. Three behaviours are evident in this figure. For short eclipse egresses, the situation closely resembles an instantaneous jump in count rate, and the bias on the position of the change point is accordingly scattered around zero, with the scatter appearing large in the figure because Δt is small. For intermediate durations, where the duration is longer than the photon separation in the post-eclipse block, the change point is triggered earlier than the half-way point of the egress. There is a bias of about 25 to 33%. Finally, when the post-eclipse count rate is high and the egress is long, then the algorithm never fails to resolve it, as indicated by the absence of such points in Fig. 4.

3. Data simulated for background subtraction

In this section we devise and apply several methods for performing background subtraction using a Bayesian Blocks approach.

Table 1. Timing accuracy of the background subtraction methods for the thirty simulated transients.

Transient	Time (s)	Δt_b	Δt_c	Δt_d	Δt_e	Δt_f	Δt_g	Δt_h
I-1	1751.0	-1.0	0.4	0.4	0.1	0.5	1.7	1.7
I-2	4786.9	3.1	-0.2	-0.2	0.4	0.1	0.4	0.4
I-3	7822.7	-2.7	4.0	4.0	-3.0	3.7	4.4	4.4
I-4	10 858.6	1.4	-0.1	-0.1	-2.3	-0.1	0.7	0.7
I-5	13 894.5	-4.5	0.1	0.1	-0.7	0.1	8.9	8.9
I-6	16 930.3	-0.3	0.1	0.1	-3.2	0.1	2.4	2.4
I-7	19 966.2	3.8	-0.2	-0.2	-0.2	0.0	6.7	6.7
I-8	23 002.1	-2.1	-8.3	68.3	-3.4	0.5	68.9	-8.2
I-9	26 037.9	2.1	-5.0	25.2	0.8	0.5	25.0	5.5
I-10	29 073.8	-3.8	1.1	1.1	-0.6	1.2	2.1	2.1
E-1	2249.0	1.0	-0.0	-0.0	0.5	0.2	-1.4	-1.4
E-2	5284.9	-4.9	0.3	0.3	0.5	-0.3	-0.2	-0.2
E-3	8320.7	-0.7	-1.3	-1.3	2.0	0.0	-1.8	-1.8
E-4	11 356.6	3.4	-0.3	-0.3	0.1	-0.0	-3.5	-3.5
E-5	14 392.5	-2.5	-0.1	-0.1	0.7	0.9	-23.2	-23.2
E-6	17 428.3	1.7	-1.4	-1.4	0.4	0.1	-2.5	-0.5
E-7	20 464.2	-4.2	-4.9	-4.9	-0.1	-4.4	-5.4	-5.4
E-8	23 500.1	-0.1	-16.4	-16.4	40.9	-1.1	-16.8	-14.1
E-9	26 535.9	4.1	16.2	44.0	35.2	-0.3	44.4	16.0
E-10	29 571.8	-1.8	-0.3	10.7	15.5	-0.7	21.5	-0.3
B-1	800.0	0.0	-0.0	-0.0	0.0	0.1	-0.3	-0.3
B-2	3345.6	4.4	-0.0	-0.0	0.1	0.1	-0.2	-0.2
B-3	5891.2	-1.2	-0.0	-0.0	0.1	0.0	-0.3	-0.3
B-4	8436.8	3.2	-0.1	-0.1	-0.6	-0.6	-0.2	-0.2
B-5	13 527.9	2.1	-0.1	-0.1	0.0	0.0	-0.4	-0.4
B-6	16 073.5	-3.5	0.0	0.0	0.2	0.2	-0.2	-0.2
B-7	18 619.1	0.9	-0.0	-0.0	0.0	0.0	-0.2	-0.2
B-8	21 164.7	-4.7	0.0	0.0	0.1	0.1	-0.0	-0.4
B-9	23 710.3	-0.3	-0.0	-0.0	0.1	0.0	-0.1	-0.1
B-10	28 801.4	-1.4	0.1	0.1	-0.9	0.1	-0.1	-0.1

Notes. The transients are grouped by type rather than chronological order, with I, E, B referring to ingresses, egresses, and bursts. The columns Δt_b through Δt_f correspond to the error in the timings: b = constant cadence, c = direct subtraction, d = weighted blocks, e = weighted cells, f = weighted photons, g = Iterated Bayesian Blocks method with original weighting, h = Iterated Bayesian Blocks method with alternate weighting.

We have tested the various methods in this paper on a hypothetical *XMM-Newton* observation of a binary system exhibiting transients in the form of eclipses and bursts. The observation of the source is affected by a soft proton flare, contributing a substantial background of photons that increases in magnitude over the observation, and is highly time-variable.

We generated two lists of photon arrival times. The first was assumed to be taken from a region surrounding the source, containing photons from the source itself and from the soft proton flare superimposed on it. The second list contains only flare photons taken from a region 4.123 times larger than the source extraction region. This number was chosen arbitrarily. Thus, photons from the background extraction region were given a weight of $W_B = 1/4.123$ compared to source region photons. We attempted to recover the source light curve by subtracting the soft proton flare from the total light curve. The ability of our background subtraction methods to recover the timings of the transient events is our measure of the suitability of the background subtraction methods. We summarize the results in Table 1 and Fig. 13.

The source and flare light curves are described in Sects. 3.1 and 3.2.

3.1. Source light curve

We took the behaviour of the LMXB EXO 0748–676 as a model for our simulated source light curve. This system exhibits

eclipses of 8.3 min duration, with an orbital period of 3.82 h (Parmar et al. 1986), during which the X-ray flux from the source drops to zero. It also undergoes Type I X-ray bursts, during which the X-ray flux increases almost instantaneously to many times its pre-burst level (Gottwald et al. 1986). When neither eclipses nor bursts are present, the source emits with a count rate of about 3 counts/s in *XMM-Newton* observations (Homan et al. 2003).

The simulated source light curve consists of a constant persistent intensity of 3 counts/s. Superimposed on this are eclipses of duration 498 s, recurring with a faster orbital period of 3035 s, and bursts recurring with a period of 2545.6 s. The bursts have peak intensity of 30 counts/s and an exponential decay time of 24 s. The transient events in the light curve are therefore (see Fig. 5) ten bursts, ten eclipse ingresses, and ten eclipse egresses. All transients occur instantaneously, and two bursts are missing because they occur during an eclipse. The times of these events are given in Table 1. There are 154 992 photons in this series.

3.2. Background light curve

The background consists of three segments. First is a linear rise from a count rate of 0 to 24, to $t = 7500$, followed by a quadratic decrease back to zero. These two segments are intended to investigate whether the presence of a rising or falling background contribution biases the timings of transients. Beginning at 15 000 s is an oscillatory signal with increasing amplitude

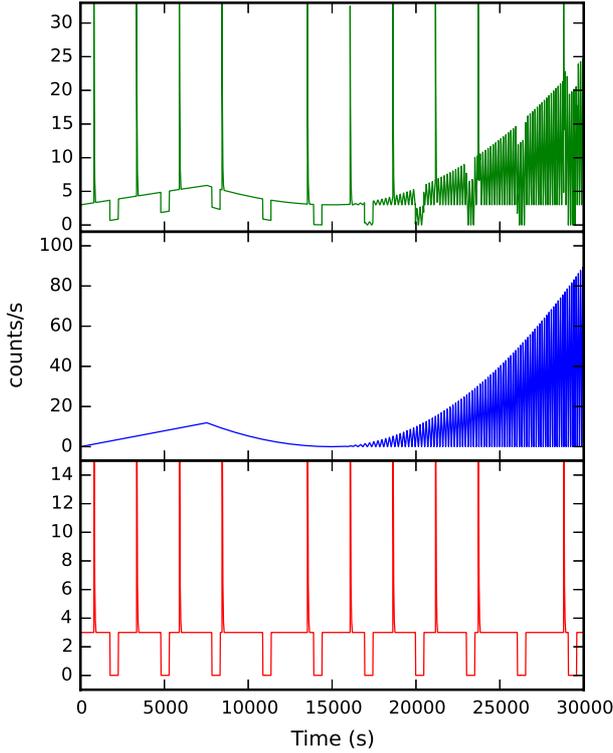


Fig. 5. Total source light curve (*top panel*), background (*middle panel*), and zero-background source light curve (*bottom panel*). The high intensity of the background signal is due to the higher exposure area (4.123 times greater than the source area).

and frequency, designed to approximate the behaviour of the soft proton flares that affect observations by X-ray instruments such as *XMM-Newton* and *Chandra* (Lumb et al. 2002). Its count rate is given by the formula

$$R(t) = 2.0 \times 10^{-7} t_a^2 \left[1 + \sin \left(\frac{1}{5000} t_a^{1.25} \log_e t_a \right) \right] \text{ photons/s,} \quad (4)$$

where $t_a = t - 15\,000$. There are 300 725 photons in this series.

4. Background subtraction

4.1. Constant cadence

Before analysing various implementations of the Bayesian Blocks algorithm, we study for comparison a case in which the locations of the change points are not optimally determined by the data. The simplest method is to group the source and background photons into equally spaced bins whose width and location are fixed beforehand, then subtracting the latter from the former to obtain the background-subtracted light curve. In Fig. 6 we show the two photon series and their difference, counted in bins of 10 s width.

A property of the constant-cadence method is that signals with more rapid variability than the cadence will tend to be averaged out. This property makes it useful for subtracting a rapidly varying background. However, fast transients such as burst rises and eclipse in- and egress (instantaneous in our simulated data) will not be timed accurately unless they fortuitously happen to occur on or near a bin boundary. An instantaneous transient event occurring well inside a bin will not appear instantaneous because it will produce one bin of intermediate count rate, giving the false impression of a gradual change; this phenomenon is

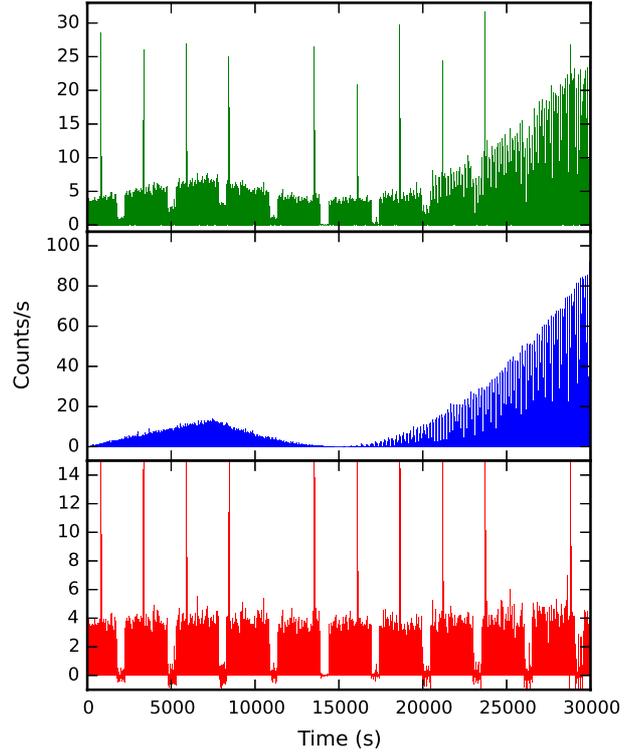


Fig. 6. Source, background, and background subtracted light curves (*top, middle and bottom panels respectively*) accumulated into bins of predetermined width (10 s) and position. This method is effective at removing the background but its ability to time transient events is limited by the previously set bin size.

clearly seen in panel b of Fig. 12. In this regime, the uncertainty in the timing of the transient event will be smaller than $\Delta t/2$, where Δt is the width of the bins.

Another disadvantage of the constant cadence-method is that it is necessary to select the width of the bins beforehand. Various estimates for the best bin width have been proposed (see Birgé & Rozenholc 2006 for a discussion), but they often suggest too few bins to perform timing analyses. The commonly used rule of Sturges (1926), taking $1 + \log_2 N$ intervals, gives only 18 or 19 bins for the 154 992 source photons. Similarly, the rules of Scott (1979) and Freedman & Diaconis (1981) give bin widths of hundreds of seconds.

4.2. Bayesian Blocks – direct subtraction

The most obvious way to perform background subtraction is to generate Bayesian Blocks light curves for the source and background regions and subtract the latter from the former, after normalizing the background series to account for the larger size of the extraction region. The results of this test are shown in Fig. 7 and panel c of Fig. 12. The resulting background-corrected light curve will contain a block boundary for every change point in both of the source and background Bayesian Blocks representations, and many of these will be spurious. Subtracting the background Bayesian Blocks light curve from the source light curve succeeds in removing much of the background. However, the resulting light curve is very noisy, particularly towards the end of the observation.

Scargle et al. (2013) provided a method for combining multiple data series in such a way that the change points in each series match. Such a process is useful for concurrent observations of

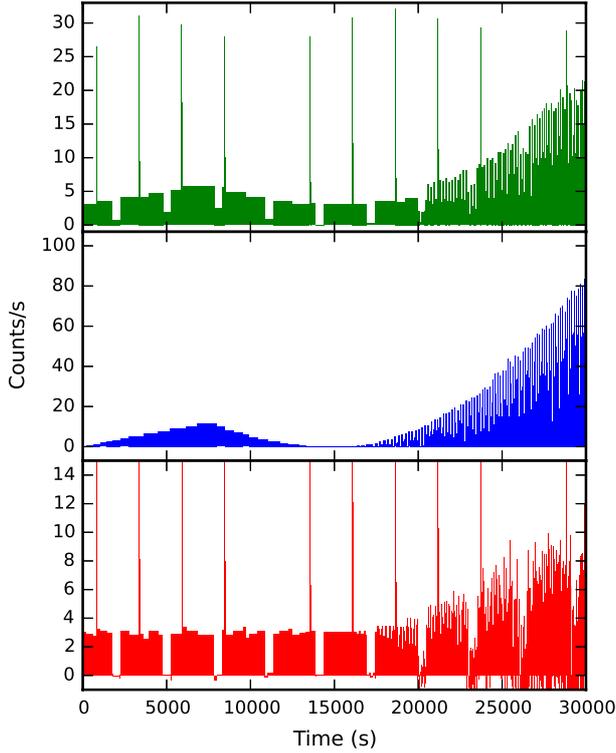


Fig. 7. Direct subtraction of the Bayesian Blocks representation of the background from that of the source. This method produces one change point for each change point in the two Bayesian Blocks representations from which it is derived.

the same transients by different instruments and will not produce superfluous change points when they are added or subtracted, but it is not suitable for data series such as ours, which contain completely different transients.

If we do not wish to produce a Bayesian Blocks representation of the background, the background can be subtracted from the source Bayesian Blocks representation on the level of blocks, cells, or individual photons. These three approaches are described in Sects. 4.3–4.5.

4.3. Bayesian Blocks – weighted blocks

This method is suggested by the observation that photons from the background region are more numerous, but individually carry less weight than source photons. Here we find the Bayesian Blocks change points of the source photon list, and subtract from the source photon count rate the exposure area-weighted count rate of the background photons falling into that block. That is,

$$\text{CR} = \frac{n_S - W_B n_B}{L}, \quad (5)$$

where CR is the background-subtracted count rate, n_S and n_B are the numbers of source-region and background-region photons, and L is the length of the block. The results are shown in Fig. 8 and in panel d of Fig. 12.

Although this method does not produce as many spurious change points as the previous one, it is clear that it cannot give the locations of the transients as accurately because it only contains the change points of the source series and these are potentially offset from their true locations by the variable background, as can clearly be seen in Fig. 12.

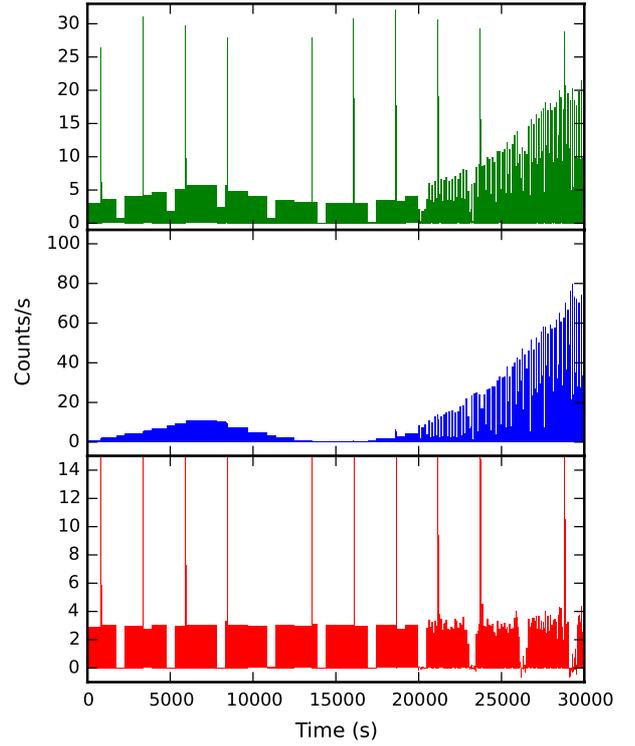


Fig. 8. Background-corrected light curve using the weighted-blocks method. The background subtraction appears to be very effective, producing almost no noise even at the end of the observation. The eighth and ninth eclipses appear to have shallowly sloping edges, probably due to pulses in the background.

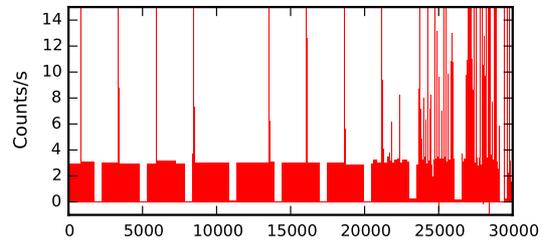


Fig. 9. Background-corrected light curve using the weighted-cells method. There is considerable noise near the end of the observation, but all the transients appear sharply defined.

4.4. Bayesian Blocks – weighted cells

In this approach we subtract the weighted background photons from each cell before the change points are found. Thus, each cell contains not one photon, but $1 - n_B W_B$ photons, where n_B is the number of background photons occurring within that cell. As the fitness function of each potential block (Eq. (19) of Scargle et al. 2013) involves the logarithm of the counts in blocks, we must find a way of dealing with cells and blocks with negative count rates. If the count rate is negative we set the effective count rate in the logarithm to a small positive number s_{\min} (we have taken $s_{\min} = 1.0 \times 10^{-4}$). The cells themselves are defined by the source photons, and therefore each contains at least one positively weighted photon, hopefully reducing the number of times we need to resort to this arbitrary countermeasure. If a block edge separates two blocks of apparently negative count rate, we did not adjust its location according to Eq. (3) because it is unclear what the spacing between two hypothetical photons in such a block should be.

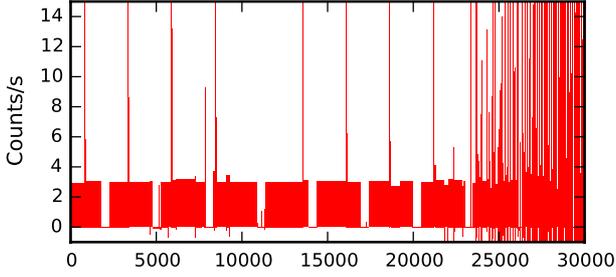


Fig. 10. Background corrected light curve using the “weighted photons” method. Noise, in the form of short blocks with implausibly high or low count rates, is more prominent in this method, but the transients are sharply defined.

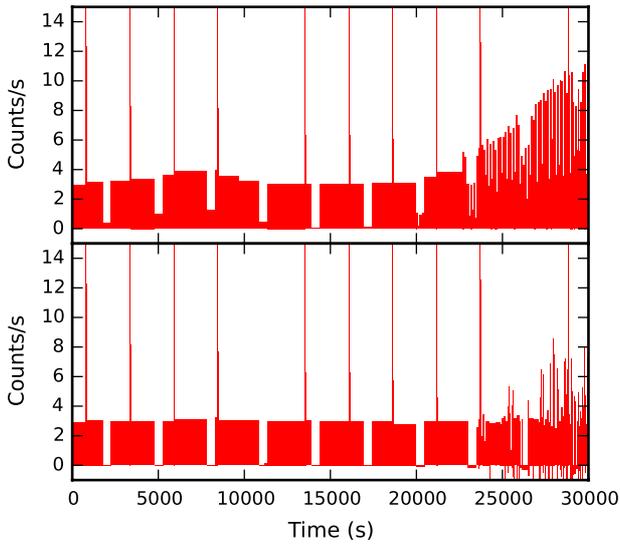


Fig. 11. Method described in Mossoux et al. (2015). *Top panel:* original photon weighting, *bottom panel:* alternative photon weighting. The original weighting does not subtract the background, but only attempts to locate the transients, which explains why the eclipses do not have count rates near zero. Both methods are only noisy near the end of the observation.

The results of this trial are shown in Fig. 9 and panel *e* of Fig. 12. The weighted-cells method performs well throughout most of the observation, but towards the end it produces many short spurious blocks with implausibly high count rates. It also misses some eclipse timings by much (see panel *e* of Fig. 12, and Table 1).

4.5. Bayesian Blocks – weighted photons

This approach is similar to the previous one, except that we no longer define the cell edges with source photons alone. Instead, we combined both photon lists into one, and each cell has a weight of either 1 or $-W_B$, depending on whether it contains a source or a background photon. As in the previous method, we did not adjust the location of the change point between two blocks with negative count rates.

The problem with negative count rates is more prominent now, because there are many cells that contain a background photon and therefore have negative count rates. However, the placement of change points is potentially finer. The results are shown in Fig. 10 and in panel *f* of Fig. 12. There are many short blocks with very high or low count rates near the end

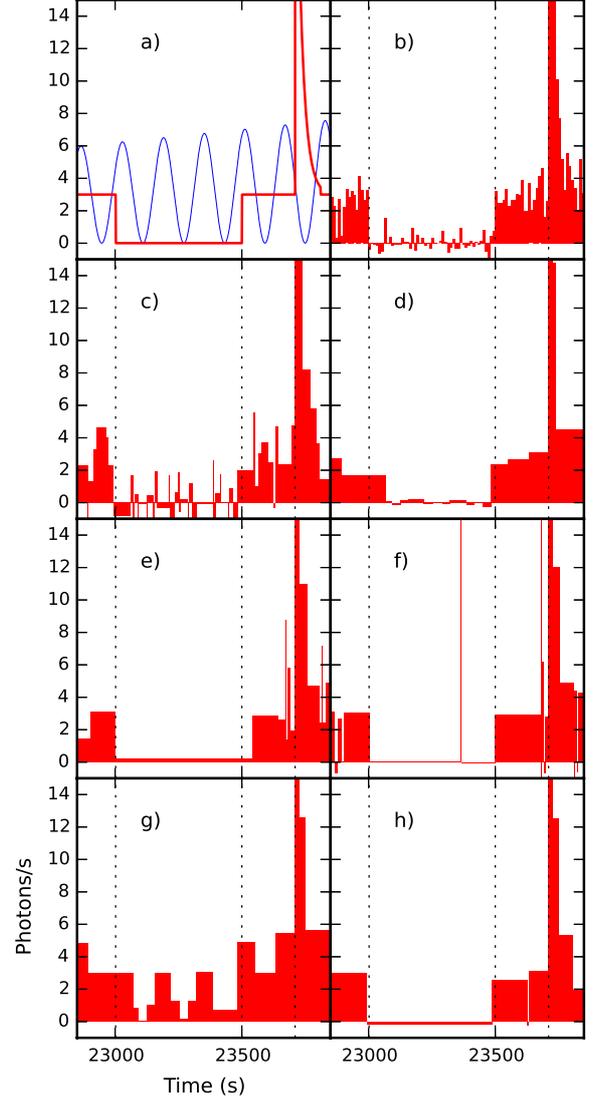


Fig. 12. Region around eclipse #8 and burst #9. **a)** true source light curve (red) and background (blue); **b)** constant cadence; **c)** direct subtraction; **d)** weighted blocks; **e)** weighted cells; **f)** weighted photons; **g)** iterated Bayesian Blocks with original weighting; **h)** iterated Bayesian Blocks with alternative weighting. The true locations of the transients are indicated by dotted vertical lines. All other methods locate the burst accurately but many fail to accurately time the eclipse ingress, which coincides with a pulse of the variable background. The weighted-photons approach performs very well in locating all transients.

of the simulated observation, and even a few nonsense blocks near the beginning. The timing properties of this approach are nonetheless excellent.

4.6. Bayesian Blocks – iterated Bayesian Blocks

This approach was developed and successfully applied to X-ray activity of Sgr A* by Mossoux et al. (2015). It is effectively a hybrid of the weighted-cells and direct-subtraction approaches. Here, a Bayesian Blocks representation is produced for the background and source light curves to obtain count rates for the source and background regions. The count rates are then used to calculate appropriate weights for photons in the source region and after which the algorithm is run a third time. A photon falling within a block of source region count rate C_S and background

region C_B is given a weight of $w = C_S/(C_S + C_B)$, where C_B is positive and corrected for exposure area. It is clear that weighting the photons in this way does not subtract the background, but is only intended to find the change point locations. For this reason, we also tested this recipe with an alternate weighting, $w = 1 - C_B/C_S$. The results are shown in Fig. 11, and in panels g and h of Fig. 12.

4.7. Invariance to time units

The unmodified Bayesian Blocks algorithm produces the same change points regardless of the choice of time unit. We have investigated whether this desirable behaviour is preserved for the various modifications described above. We took the same photon series and divided the arrival times by 3600 to express the arrival times in hours rather than seconds. Naturally, the constant cadence, direct subtraction, and weighted blocks methods were unchanged under this transformation, but all other methods produced many more change points than previously. The weighted-photons method, for example, produced 16 080 change points for the time-scaled photon series, compared to 971 for the original series. This behaviour is probably due to our safeguard against negative block count rates. Since we have achieved tolerably good results with count rates of 1 to 10, but higher effective count rates produce very many spurious change points, it may be generally desirable to scale the time unit. We have also found that the number of change points produced is fairly insensitive to the choice of s_{\min} .

It would be desirable to have a method for dealing with blocks of negative count rate that has a better theoretical justification. Similar questions have previously been considered. Loredo (1992)¹ gives a posterior probability distribution for the source count rate, independent of the count rate of the background, which might be used to ensure positive count rates in all potential blocks. A method developed by Zech (1989) gives the probability of the source having a positive count rate s given the number of total (source + background) photons detected, and the background count rate. This can be used to estimate the highest plausible source count rate. However, it is not immediately obvious how to incorporate either of these procedures into the fitness function of Eq. (19) in Scargle et al. (2013). Both procedures also require summations over all the photons in a block, increasing the running time of the algorithm from $O(n^2)$ to $O(n^3)$ or longer.

4.8. Background exposure area

We investigated the effect of changing the exposure area of the background by repeating the weighted-photons method on two new simulated observations, using background-extraction areas of 2.062 and 8.246 as large as the source extraction region, that is, half and twice the exposure area we have been using up to now. The background photons were weighted according to these new exposure areas.

For the three background exposure areas, we found that the smallest one produced 1302 change points, the middle one produced 971 change points, and the largest background exposure area produced 742. The accuracy in timing the thirty transient events was not adversely affected by increasing or decreasing

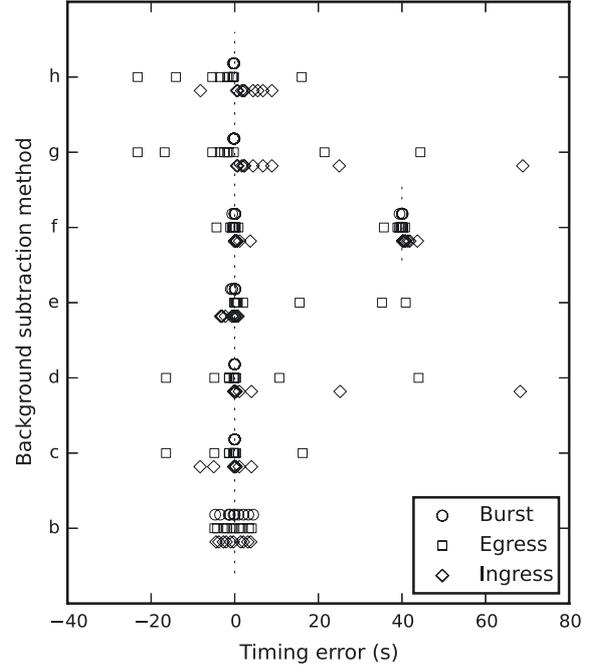


Fig. 13. Timing errors of the thirty transient events, for the different background-subtraction methods. The letter identifiers are the same as in Table 1 and Fig. 12: **b)** constant cadence; **c)** direct subtraction, **d)** weighted blocks; **e)** weighted cells; **f)** weighted photons; **g)** iterated Bayesian Blocks with original weighting; **h)** iterated Bayesian Blocks with alternative weighting. The transient events have been divided by type. The constant-cadence method finds all types of transients with precision limited by the bin size. All other methods perform well for bursts, but only the weighted-photons approach **f)** is reliably accurate in timing eclipse ingresses and egresses. The offset points in row **f)** show the timing errors for this method if the locations of the change points are not adjusted according to Eq. (3). Dotted vertical lines indicate a zero timing error.

the background exposure area, but the larger this area, the fewer spurious change points were produced.

5. Discussion and conclusion

We have investigated the ability of the Bayesian Blocks algorithm to accurately and precisely time transient events. With an adjustment that moves the change point toward the block of higher count rate, the algorithm is able to determine the locations of instantaneous increases or decreases in count rate with essentially no systematic bias and with uncertainties similar in size to the interval between individual photons.

When the change in count rate was not instantaneous, we characterised the ability of the algorithm to resolve the period of varying count rate and, if it could not be resolved, found that the change point was placed near the block of lower count rate. These observations have important implications for applications such as measuring the times and durations of eclipses.

We have incorporated background subtraction into the Bayesian Blocks algorithm in several different ways, and tested the alternatives against simulated source and background signals containing bursts and eclipses of known position. Even when the observation was dominated by a strong, rapidly varying, background contamination it was possible to recover the transient times with excellent accuracy. In Table 1 we show the time locations and nature (e.g., I-3 is the third eclipse ingress) of the transients, as well as the error (the distance to the nearest bin edge

¹ A longer and more comprehensive version of this book chapter is available at <http://www.astro.cornell.edu/staff/loredo/bayes/tj1.html>

found by each of the background-subtraction methods). The results are also shown graphically in Fig. 13. Note that, because we knew beforehand where the transients were, we were able to unambiguously identify the nearest bin edge. In real applications it may be difficult to determine which bin edge marks the beginning of the transient, therefore methods producing many needless bin edges or change points may perform less well on real data than in this paper. Visual inspection of the results is always called for.

The weighted-photons approach (Sect. 4.5) is clearly best suited for background subtraction. As shown in Table 1, it correctly recovers all transients to within a few seconds and most to within half a second. However, it is prone to producing occasional blocks of very short lengths and implausibly high count rates (see Figs. 10 and 12), but if this drawback can be tolerated then the weighted-photons adaption to the Bayesian Blocks algorithm is well suited for transient timing. All alternatives perform well in locating the bursts, but only the weighted-photons approach reliably recovers the eclipse ingresses and egresses. The reason for this can be seen in Fig. 12: a pulse of background contamination coincides with an eclipse ingress, which causes the weighted-blocks and unmodified iterated Bayesian Blocks methods to miss the ingress entirely, and most of the other alternatives can recover it only approximately.

The effect of adjusting the change points according to Eq. (3) is also shown in Fig. 13. The offset points for the weighted-photons row show the effect of not performing this adjustment. There is little visible difference, but the timings of some of the ingresses are visibly improved by performing the change point adjustment.

We included a slowly rising and falling background in the simulated observation to determine whether this causes a systematic offset in the timings but according to Table 1, there is no indication of such an effect in any of the alternatives we considered.

The modifications to the Bayesian Blocks algorithms developed in this paper can be generalised to be applied to simultaneous observations with different instruments, each with their own source and background-extraction regions. Every photon from all data series is assigned a weight according to the instrument's extraction area and sensitivity.

The unmodified Bayesian Blocks algorithm is invariant for different choices of time unit. That is, it makes no difference to the number and placement of the change points if the photon arrival times are expressed in units of seconds, hours, orbital phase, etc. We have found that this useful property is not preserved for modifications to the Bayesian Blocks algorithm in which cells or blocks can have negative weight, probably because of the way we avoided taking the logarithm of a negative photon count. The higher the count rate, the more spurious change points will be produced. If the unit of time is scaled to achieve typical count rates of 1–10, the results are expected to still be acceptable.

Acknowledgements. This work was supported by the German DLR under contract 50 OR 1405. This research made use of Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration, [Astropy Collaboration et al. 2013](#)). We are grateful to the referee, whose suggestions significantly improved this work.

References

- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, **558**, [A33](#)
- Bélangier, G. 2013, *ApJ*, **773**, [66](#)
- Birgé, L., & Rozenholc, Y. 2006, *European Series in Applied and Industrial Mathematics: Probability and Statistics*, **10**, 24
- Burgess, J. M. 2014, *MNRAS*, **445**, [2589](#)
- Freedman, D., & Diaconis, P. 1981, *Z. Wahrscheinlichkeitstheorie Verwandte Gebiete*, **57**, [453](#)
- Gottwald, M., Haberl, F., Parmar, A. N., & White, N. E. 1986, *ApJ*, **308**, [213](#)
- Homan, J., Wijnands, R., & van den Berg, M. 2003, *A&A*, **412**, [799](#)
- Ivezić, Ž., Connolly, A., Vanderplas, J., & Gray, A. 2014, *Statistics, Data Mining and Machine Learning in Astronomy* (Princeton University Press)
- Knuth, K. H. 2006 [[arXiv:physics/0605197](#)]
- Loredo, T. J. 1992, in *Statistical Challenges in Mod. Astron.*, eds. E. D. Feigelson, & G. J. Babu, [275](#)
- Lumb, D. H., Warwick, R. S., Page, M., & De Luca, A. 2002, *A&A*, **389**, [93](#)
- Mossoux, E., Grosso, N., Vincent, F. H., & Porquet, D. 2015, *A&A*, **573**, [A46](#)
- Parmar, A. N., White, N. E., Giommi, P., & Gottwald, M. 1986, *ApJ*, **308**, [199](#)
- Scargle, J. D. 1998, *ApJ*, **504**, [405](#)
- Scargle, J. D., Norris, J. P., Jackson, B., & Chiang, J. 2013, *ApJ*, **764**, [167](#)
- Scott, D. W. 1979, *Biometrika*, **66**, [605](#)
- Stelzer, B., Flaccomio, E., Briggs, K., et al. 2007, *A&A*, **468**, [463](#)
- Sturges, H. A. 1926, *J. Am. Stat. Assoc.*, **21**, [65](#)
- Zech, G. 1989, *Nucl. Instrum. Meth. Phys. Res. A*, **277**, [608](#)