J. Astrophys. Astr. (2021) 42:73
https://doi.org/10.1007/s12036-021-09714-6
© Indian Academy of Sciences

SCIENCE RESULTS

The search for fast transients with CZTI

Y. SHARMA[1,2], A. MARATHE[2,3], V. BHALERAO[2,*], V. SHENOY[2], G. WARATKAR[2], D. NADELLA[3], P. PAGE[2], P. HEBBAR[2,4], A. VIBHUTE[5], D. BHATTACHARYA[5], A. R. RAO[6] and S. VADAWALE[7]

[1]Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA.
[2]Indian Institute of Technology Bombay, Powai, Mumbai 400 076, India.
[3]National Institute of Technology Karnataka, Surathkal, Mangalore 575 025, India.
[4] University of Alberta, Edmonton, AB T6G 2E1, Canada.
[5]Inter-University Centre for Astronomy and Astrophysics, P.O. Bag 4, Ganeshkhind, Pune 411 007, India.
[6]Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400 005, India.
[7]Physical Research Laboratory, Ahmedabad 380 009, India.
*Corresponding Author. E-mail: varunb@iitb.ac.in

MS received 3 November 2020; accepted 17 January 2021

Abstract. The Cadmium–Zinc–Telluride Imager on AstroSat has proven to be a very effective All-Sky monitor in the hard X-ray regime, detecting over three hundred GRBs and putting highly competitive upper limits on X-ray emissions from gravitational wave sources and fast radio bursts. We present the algorithms used for searching for such transient sources in CZTI data, and for calculating upper limits in case of non-detections. We introduce CIFT: the CZTI Interface for Fast Transients, a framework used to streamline these processes. We present details of 87 new GRBs detected by this framework that were previously not detected in CZTI.

Keywords. Stars: gamma-ray burst: general—X-rays: bursts—methods: data analysis.

1. Introduction

The Cadmium–Zinc–Telluride Imager (CZTI, Bhalerao *et al.* 2017b) is a high-energy coded aperture mask instrument on board AstroSat (Singh *et al.* 2014) CZTI comprises of four independent, identical quadrants giving a total physical area of 976 cm^2. Each quadrant consists of a 4×4 array of 5-mm thick Cadmium–Zinc–Telluride detectors, giving good sensitivity in the 20–200 keV energy range and an energy resolution of 11% at 60 keV. In nominal operations, all incident photons are saved in event-mode with 20 μs resolution. While primary coded field-of-view of CZTI is $4.6^\circ \times 4.6^\circ$, the collimators and support structure of CZTI become increasingly transparent to radiation at energies above ~ 100 keV, making it sensitive to sources all over the sky. The off-axis sensitivity depends on the effective area, which in turn is a strong function of energy and direction. Details of the effective area calculations are presented in Mate *et al.* (2021). As there are very few bright sources in this energy range, the net contribution of off-axis sources is small and simply manifests itself as a slightly elevated background.

A special exception to this are bright, short-duration transient sources like gamma ray bursts (GRBs). GRBs with their high brightness and short durations (seconds to minutes) manifest themselves as an increase in the count rates in CZTI. Starting from the first GRB detection on the day the instrument was powered on (GRB 151006A; Bhalerao *et al.* 2015; Rao *et al.* 2016), CZTI has detected 325 GRBs in the five years since launch. On the other hand, the lack of a measurable change in count rates

This article is part of the Special Issuse on "AstroSat: Five Years in Orbit".

Published online: 06 July 2021

corresponding to a transient event can be mapped to an upper limit on the flux of the transient. With this technique, we have obtained stringent upper limits on X-ray emission from Fast Radio Bursts (Anumarlapudi *et al.* 2020), as well as from gravitational wave sources (Bhalerao *et al.* 2017a).

In this paper, we describe the methods used for searching for such sources (called fast transients hereafter). In Section 2, we discuss the pre-processing of data for our searches. In Section 3, we discuss the search for "known" transients, where the time and possibly location are known from other sources. We also discuss methods for putting upper limits on the flux from such transients in case they are not detected in data. In Section 4, we discuss in detail the algorithms, software, and the interface developed for searching for transients in all of CZTI data. In Section 5, we discuss the performance of our software, and present the 87 transients detected in our searches. We conclude by discussing future improvements in Section 6.

2. Preparing the data

The CZTI data reduction pipeline[1] is designed for imaging and spectroscopy of sources in the primary field of view. There are two particular operations in the pipeline that are detrimental to the search and analysis of fast transients. First, the pipeline discards data from time intervals when the on-axis source being targeted by AstroSat is occulted behind earth — though CZTI might still detect fast transients that are located elsewhere in the sky. Second, sections of data where the count rates in detectors rise above a certain value are discarded as noisy: thus suppressing bright transients. For fast transient searches, we overcome these issues by changing a few pipeline parameters — thus ensuring that final data products are still compatible with any post-processing software. We follow the standard procedure to obtain Level-2 "bunch cleaned" data created by `cztbunchclean`. Next, when selecting good time intervals with `cztgtigen`, we change the config file `mkfThresholds.txt` to remove the earth occult condition (the ELV parameter), which would have discarded data when the on-axis target was occulted by the earth. The next stage is to reject noisy sections of data using `cztpixclean`. The default settings of `cztpixclean` discard intervals where a single pixel has more than 2 counts per second, or where a module has more than 35 counts per second. To ensure that this step does not discard bright transients, we raise the detector count threshold to 1000 and the pixel count threshold to 100. Finally, we run `cztevtclean` to obtain cleaned event files. Since our processing is done independently for each quadrant, we use the `_quad_clean.evt` files.

The next stage is to create light curves for each quadrant. Here we have to carefully correct for various sources of dead time in the instrument: for instance quadrant-level dead time (0.3 s dead time for collecting housekeeping data every 100 s), and module-wise dead time (arising from discarding particle–induced photon bunches). We use the pipeline module `cztbindata` to consider all these factors to correctly calculate the dead time for each time bin used. For certain searches, we also limit select the photon energy ranges in this step.

The final step in data preparation is to remove the orbit-induced trends in the background. As AstroSat is in low earth orbit, the satellite sees a variable background count rate over different parts of the earth, rising near the South Atlantic Anomaly (SAA). We see that the background variations are relatively smooth, over timescales of hundreds of seconds. But, if a transient event were to evolve on comparable or longer timescales, we would not be able to distinguish it from background variations. Fortuitously, most transients of interest have timescales of tens of seconds or shorter. Hence, we can fit a smooth trend to the data and subtract it, effectively making the data "background-free" and greatly simplifying the task of transient detection. We have tested two methods for de-trending the data: in the first method, the trend is estimated by using a running median filter of 100 second width. In the second method, we estimate the background using a second order Savitzky-Golay (savgol) filter of 100 second width (for details see Anumarlapudi *et al.* 2020). Both trend estimates work well, and hence both are coded into our software. In preliminary testing, the savgol filter yielded better results for transient searches, hence it is set as the default filter.

3. Triggered searches

In CZTI data analysis, searches for fast transients are broadly categorised into two types: "triggered" and "blind". Triggered searches are cases where the time of a transient, and possibly its position, are already known. For such cases, a qualitative search is carried out by pre-processing the data followed by visual

[1]CZTI pipeline: http://astrosat-ssc.iucaa.in/?q=cztiData.

examination. Blind searches, that are more quantitative, are discussed in Section 4.

3.1 *Method*

Triggered searches start with pre-processing the data as discussed in Section 2, up to the creation of cleaned event files. We then create "spectrograms" or "time-energy plots": two-dimensional histograms of the event data, and visually examine them for the transient (Fig. 1(a)). By default, the energy axis is binned in 10 keV bins from 20–200 keV. Searches are carried out by binning the time axis in 0.1 s, 1 s, and 10 s bins. We also calculate two further variants of this spectrogram to aid visual searches: we calculate the mean spectrum and subtract it from each time bin, thus highlighting any transient variations (Fig. 1(b)). In the third step, we take these mean-subtracted spectrograms and normalise the light curve in each energy bin by its standard deviation (Fig. 1(c)). This de-weights noisy energy bands, and gives a rough idea of the statistical significance of any transient.

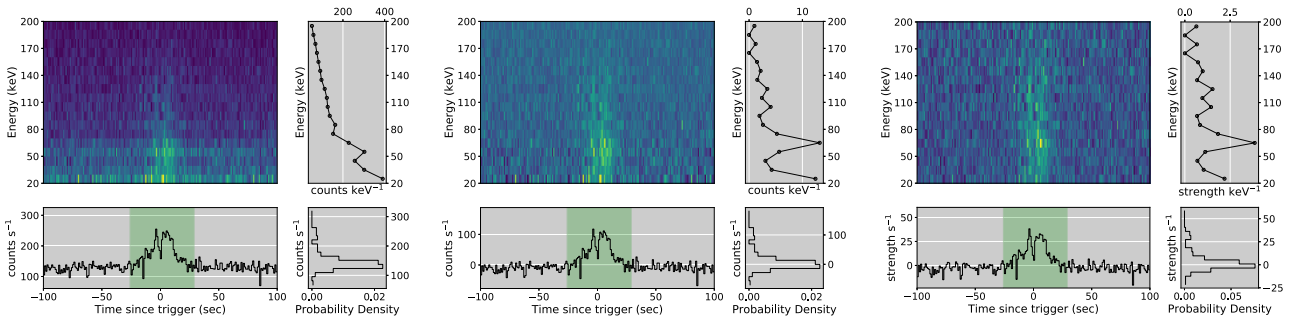
Light curves from a single quadrant occasionally show noise spikes which look similar to astrophysical transients. These events — often caused by charged particles or electronic noise — typically occur at low energies ($\lesssim$50 keV). Since the four quadrants of CZTI are electronically independent, the electronic noise events are always caused in just a single quadrant. Such noise candidates are readily rejected by requiring that any transient is considered "detected" only if it is detected across multiple energy bins, and seen in more than one of the four independent quadrants of CZTI. Track-like events created by charged particles can sometimes be simultaneously seen in multiple quadrants. Such cases are always of short duration (<1 s), and can be discarded based on their track-like count distributions in the detector plane. Overall, four quadrant detections of transients are most unambiguous, but detections coincident in three or two quadrants are also considered acceptable if they pass the above cuts, are bright and broadband.

CZTI also has caesium iodide scintillators as anti-coincidence "Veto" detectors, to reject particle events. Veto detector spectra are sampled once per second, and downlinked along with CZT data. We generate similar spectrograms and light curves for Veto data and repeat the transient search. Since data are intrinsically binned at 1 s, the default searches are carried out only at 1 s and 10 s timescales.

These searches are typically run by the Payload Operations Centre (POC) at IUCAA. Transients detected thus are reported in GCN circulars (see for instance, Gupta *et al.* 2020; Bhalerao *et al.* 2016) and announced on the CZTI GRB page at http://astrosat.iucaa.in/czti/?q=grb, along with the associated spectrograms.

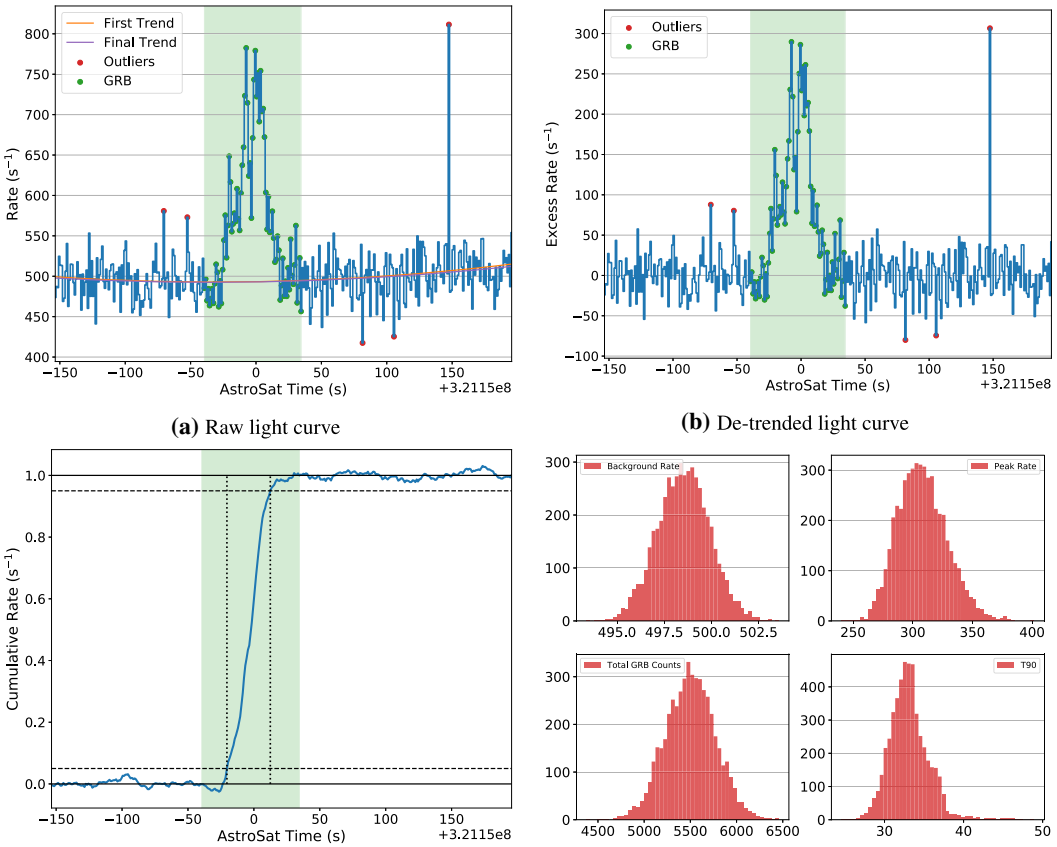
3.2 *Transient properties*

For every detected transient, we estimate its duration (T_{90}), peak rate (R_p) above background (R_b), and the total counts ($C_{\rm tot}$). We create a combined 20–200 keV



(a) Raw spectrogram **(b)** Mean-subtracted spectrogram **(c)** Normalised spectrogram

Figure 1. Spectrograms for quadrant C data for GRB 200306A, utilised in visual inspection of transient candidates (Section 3.1). Panel (**a**): The upper left frame shows raw data, binned in 1 s and 10 keV bins along the X and Y axes respectively. The upper right frame shows the spectrum, obtained by summing the spectrogram along the X axis. The lower left frame shows the light curve, obtained by summing the spectrogram along the Y axis. The lower right frame shows the distribution of count rates in the light curve. Panel (**b**): Mean-subtracted spectrogram, obtained by subtracting the average spectrum from each time bin. The four frames are analogous to Panel (a). Panel (**c**): Mean subtracted and sigma-normalised spectrogram. Note that the transient is brightest at the lowest energy bins (Panel (a)), but since those energies also have a higher sigma, the transient is statistically most significant around 60 keV (Panel (c)).



(a) Raw light curve

(b) De-trended light curve

(c) Normalised cumulative count rate

(d) Distributions of calculated parameters

Figure 2. Calculation of transient properties, illustrated with the light curve of GRB 200306A. Panel (**a**): The raw 20–200 keV light curve summed across four quadrants, with the transient region marked in green. An initial background trend (orange) is fit to the background outside the transient region, and refined (purple) with sigma-clipping outlier rejection. Outliers are marked with red circles. The mean value of the refined trend is reported as the background count rate, R_b. Panel (**b**): De-trended light curve obtained by subtracting the background trend. The peak count rate (R_p) and total counts ($C_{\rm tot}$) are measured from this de-trended light curve. Panel (**c**): A cumulative light curve calculated from Panel (b), normalised such that the median pre- and post-transient values are 0 and 1 respectively. Dashed lines indicate the points where data cross the 5% and 95% levels, which is used to calculate T_{90}. Panel (**d**): Multiple light curves are generated from Panel (a) by assuming Poisson noise distribution, and the four parameters are measured for each of these. The four frames, clockwise from upper left, show distributions of R_b, R_p, T_{90}, and $C_{\rm tot}$ obtained from these light curves. These distributions are used to define 90% confidence error bars for the parameters actually measured in Panels (b) and (c). For GRB 200306A, we get $R_b = 495^{+4}_{-3}$ counts s^{-1}, $R_p = 289^{+51}_{-19}$ counts s^{-1}, $T_{90} = 32^{+4}_{-7}$ s, and $C_{\rm tot} = 5444^{+449}_{-1023}$ counts.

light curve from all quadrants that show a clear detection of the transient. "Pre-transient" and "post-transient" sections of the light curve are visually identified, and the background is estimated by fitting a quadratic to these. The best-fit quadratic is subtracted from the data to obtain a background-free light curve, and counts are summed to create a cumulative light curve. The post-transient part of this curve gives a measure of the total counts in the transient. The time taken for the cumulative curve to rise from 5% to 95% of the total counts is the T_{90} duration of the transient (Fig. 2). These details are included in the published GCN circulars.

We use a Monte-Carlo approach to estimate the uncertainties in the GRB parameters. We assume that observed photons follow a Poisson distribution, and for simplicity use the observed number of photons in each bin as the rate (λ) parameter for the Poisson distribution in that bin. We create 5000 simulated light curves by drawing photons from such Poisson distributions for each bin, and measure the four parameters T_{90}, R_p, R_b and $C_{\rm tot}$ for each simulated light curve. We use 5–95% range in the histograms of these parameters (Fig. 2(d)) as the 90% credible intervals. For instance, the observed light curve of GRB 200306A

yields $T_{90} = 32$ s (Fig. 2(c)), while the central 90% credible region is from 25 s to 36 s (Fig. 2(d)). We report this as $T_{90} = 32^{+4}_{-7}$ s.

3.3 *Count rate limits for non-detections*

In cases where no transient is seen, we can place upper limits on the maximum counts received from the transient that would be consistent with noise. Since the mean background level varies through the orbit, we cannot use a direct rate. Instead, we de-trend the data as discussed in Section 2. In addition, due to the noise spikes discussed in Section 3.1, the distribution of count rates deviates significantly from a simple Poisson or Normal distribution. In particular, there is a large tail of positive counts with respect to the mean rate which can mimic transient signals. To overcome this hurdle of an unmodeled count rate distribution, we estimate the upper limits (hereafter referred to as cutoff rates) using data from nearby orbits. The method is based on the assumption that the rate of astrophysical transients detectable by CZTI is low enough that nearby orbits are unlikely to have a large number of transients.

We first decide the width of the window used for transient search, say $t_w = 100$ s, and an acceptable false positives probability (FPP, $\mathcal{F}$). We typically set $\mathcal{F} = 0.1$ for a single quadrant. Since we place limits using data from all four independent quadrants, the combined FPP is 10^{-4}. We now need to find a "cut-off rate" R_c such that the probability of this threshold being crossed by chance in t_w is $\mathcal{F}$. To calculate R_c, we select five orbits before and after the transient (excluding the orbit containing the transient) as "witness" orbits. We create light curves for these orbits using the same time bin as used in the original analysis, then de-trend them, and create histograms of the de-trended counts. R_c is defined as the point such that a fraction $\mathcal{F}$ of the data points have counts $> R_c$. A typical orbit has 4000–5000 s of usable data, so that analysis of ten orbits with parameters $\mathcal{F} = 0.1$ and $t_w = 100$ s ensure that 40–50 data points are above R_c. This makes the method robust to the presence of another transient in the witness orbits.

There are some caveats to be noted here. Occasionally, a quadrant can be extremely noisy in some orbit. If the candidate transient is in such an orbit, that quadrant is excluded from further analysis and there is a corresponding decrease in the FPP (for instance, Mate *et al.* 2017; Marathe *et al.* 2019). Our FPP estimates are derived from the probability of getting counts $> R_c$ in each of the four quadrants anywhere in the t_wFor all three bands, we process the window. In practice, we consider something a detection only if such spikes in counts are coincident across multiple quadrants, hence the actual FPP is even lower.

3.4 *Flux calculations*

Incident photons from off-axis transients are heavily re-processed (scattering, absorption, fluorescence, etc) by various satellite elements before they are incident on the detector. Hence, the mapping of incident spectra to measured spectra must be done by simulating these effects in software. We accomplish this by using a GEANT4-based mass model of the entire satellite (Mate *et al.* 2021). Since the effect of the satellite varies with direction, the simulations require knowledge of the source position in satellite coordinates. For transients where the position is known, Chattopadhyay *et al.* (2021) discuss a method of estimating the source spectrum and flux from CZTI data.

While methods for calculating the source spectrum are still under development, we have found that source flux calculations based on the mass model are quite reliable if the source spectrum is known from other instruments. We leverage this by assuming a power-law or band model spectrum for sources, and calculating the flux corresponding to the number of counts in a quadrant. The total flux from the source is the sum of fluxes in all four quadrants.

For certain transients, most notably gravitational wave events, the source location is not known precisely. Instead, discovery teams provide a sky-map with the source position probability distribution. For such sources, we evaluate the flux limit at each point on the sky map that is not occulted by the Earth at the instant of the transient. The overall flux limit is evaluated as a probability-weighted mean of these values (for instance, see Shenoy *et al.* 2020).

4. Blind searches for transients

The triggered searches are complemented by a broad "blind" search over all of CZTI data to identify astrophysical transients. We have two pipelines for such searches — a pipeline based on Machine Learning (ML) (Abraham *et al.* 2019) and the CIFT.[2] In this section, we discuss CIFT in detail.

[2]CIFT is pronounced as sift.

The broad outline for the CIFT searches is as follows: First, data are reduced and de-trended as discussed in Section 2. Various algorithms are used to identify outliers in light curves. These outliers are used to create 'peak maps' to identify candidate transients in data. Flagged candidates are displayed on an interface for human vetting. They undergo similar quality checks and inspection as discussed in Section 3, and final selected transients are saved in a database.

4.1 *Preparing the data*

CZTI Level 2 bunch cleaned files are organised into 'Obs-ID's which have all the data taken during observations of any particular object requested by an observer. We undertake most of our searches Obs-ID wise, thus typically processing a few to a dozen orbits at a time. We see that noise events are more frequent in lower energies, while data are cleanest at higher energies. To leverage this factor, we divide CZTI data into three energy bands: 20–50 keV, for 50–100 keV, and 100–200 keV. For all three bands, we process the data following steps from Section 2, and create de-trended light curves with 0.1 s, 1 s, and 10 s bins. We also use a 0.01 s binning when searching for counterparts to fast radio bursts. We use the entire energy range for the Veto detector, and create light curves at 1 s and 10 s binning.

Thus, we generally create 36 light curves for CZTI data (3 time bins $\times$ 3 energy bands $\times$ 4 quadrants) and 8 light curves for Veto data (2 time bins $\times$ 4 quadrants) per Obs-ID. We run a search algorithm on each light curve to identify outliers and create 'peak maps': boolean masks with value 1 for time bins containing the outliers, and 0 elsewhere. The twelve CZTI peak maps are added together, and any bin with a mask value of four or higher is flagged as a candidate transient. Similarly, the four Veto masks are combined and bins with mask value ≥ 3 are flagged as candidate transients. Next, we discuss the three outlier search algorithms currently implemented in CIFT.

4.2 *Top-N*

The Top-N (TN) algorithm is based on a simple heuristic: a transient is expected to have among the highest count rates seen in a given light curve. We identify the brightest N bins in a light curve and flag them as outliers for the peak map.

While testing this algorithm, we obtained better results if the searches were carried out one orbit at a time (as opposed to Obs-ID wise searches for other algorithms). By varying values of N, we obtained the best results for $N = 3$.

4.3 *N-sigma*

The N-Sigma (NS) algorithm is a straightforward statistics-based method to select outliers in a time series. We identify outliers by using iterative sigma clipping as implemented in the `Astropy sigma_clipped_stats` module. Starting with a de-trended light curve, we calculate the median and standard deviation (σ) values, and reject outliers that deviate more than 3σ from the median. The process is repeated with the new light curve until convergence is attained, subject to a maximum cap of five iterations. The mean value μ and the standard deviation σ of the final iteration become the key parameters of algorithm. Using these values, outliers are defined as data points with counts $> \mu + N\sigma$, where our default value is $N = 5$. The typical thresholds for flagging these outliers for various time bins, energy bands, and both detector types are given in Table 1. These values were calculated from data of entire five years of the search. We reiterate that the namesake N of this method is used only in identifying outliers for the peak map, while the iterative sigma estimation is always done at a three-sigma level.

4.4 *Cutoff rates based on false positive probability*

The cutoff rate based search (CR) algorithm aims at attaining a given False Positive Probability (FPP) for candidate transients. Cutoff rates are determined following the procedure discussed in Section 3.3, with one important distinction. In Section 3.3, we assumed the presence of transient-free data of an order of magnitude larger duration than the timespan of interest. Since CIFT searches are meant to be conducted over all available data, this requirement clearly cannot be met. Instead, we set our FPP threshold based on the expected rates of transients, in particular, GRBs.

The rate of detectable GRBs is a function of instrument sensitivity, energy range, and field-of-view. As a baseline, we note that on average *Fermi* GBM detects a GRB every 1.5 days (von Kienlin *et al.* 2020), while the BAT on the *Neil Gehrels Swift*

Table 1. Combined cut-offs for cutoff rate and NSigma methods for each binning and band. These representative rates were calculated by using all the the five years of data used in this study. Note that rates are in units of counts/s, not counts per bin.

		Band-wise cutoff		
Method	Binning (s)	0	1	2
CZTI				
Cutoff rate	0.1	1954	488	428
	1.0	263	107	102
	10.0	28	22	22
NSigma	0.1	396	410	407
	1.0	137	133	131
	10.0	41	38	38
Method	Binning (s)	Combined cutoffs		
Veto				
Cutoff rate	1.0	319		
	10.0	690		
NSigma	1.0	394		
	10.0	1150		

Observatory averages one GRB every four days (Lien *et al.* 2016). Based on these we stipulate a rough upper bound of the rate of GRBs detectable by CZTI as 0.5 GRBs per day.[3] We then stipulate that only 1% of our GRBs may be false positives (FPP = 0.01), corresponding to one false positive every 200 days.

To arrive at an approximate solution for the FPP criterion, we consider the case of searching for a GRB with 1 s duration in light curves with 1 s binning. In this scenario, our false positive requirement of 1 per 200 days maps to one false positive in 1.728×10^7 bins. Since most basic acceptance criterion is coincident detection in two or more independent quadrants, each quadrant can have one false positive in $\sqrt{17280000}$ time bins, or 4156 s. This is a significant fraction of an orbit, and hence the robust estimation of R_c requires data from several orbits. Decreasing the time bin size increases the number of samples in the light curve, and owing to the random underlying process, makes outliers more likely. To correct for this, we change our cutoff rate requirements based on the bin size $t_{\rm bin}$: R_c is selected such that a fraction $0.01 \times (t_{\rm bin}/4156{\rm s})$ of bins have a count rate $> R_c$.

We note that this is a highly simplified argument, which ignores the 12 light curves we make for every time bin and the > 4 peak map condition. It also ignores the small effect of presence of transients in our "witness" data sets. However, it serves as a good approximate argument for selecting our R_c thresholds from data.

The typical thresholds for flagging these outliers for various time bins, energy bands, and both detector types are given in Table 1. As in the NS method, these representative rates shown in the table were calculated with all five years included in this work. Some specialised searches use the entire 20–200 keV range as a single band. For such searches with 1 s binning, the cutoff rates for the 4 quadrants are 79, 68, 68, and 69 counts/sec respectively. For searches with 10 s binning, the rates drop to 10, 10, 10, and 12 counts/s respectively, corresponding to a total of 420 counts per 10 s bin.

4.5 *The CIFT interface*

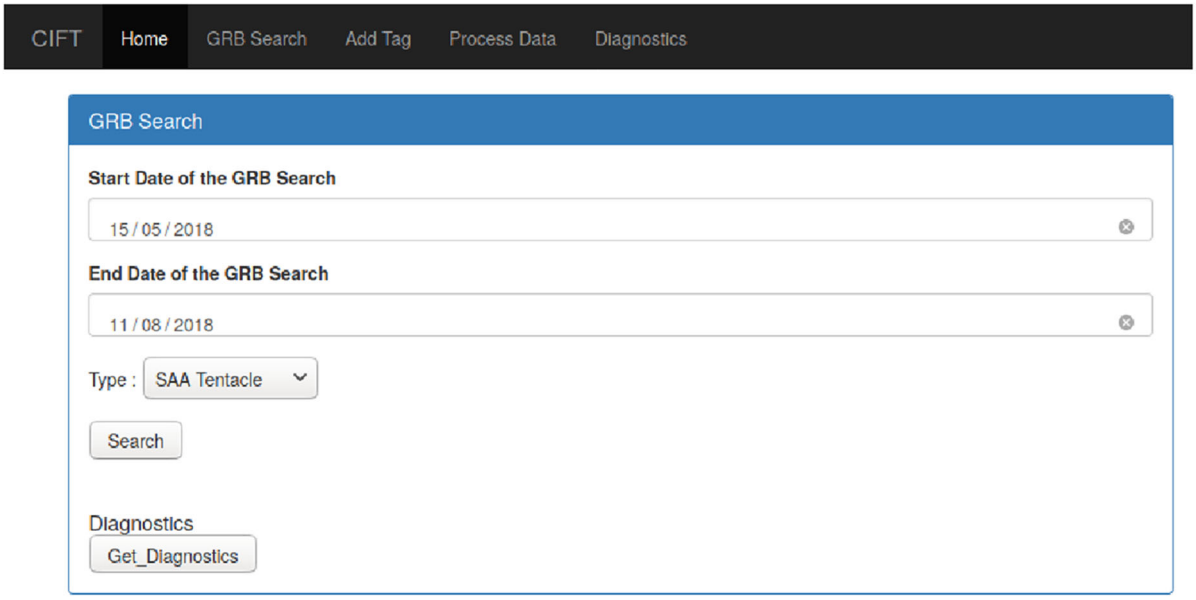
Once the peak maps have been created by any of the three algorithms discussed above, we apply our candidate selection criteria of requiring ≥ 4 matches out of 12 light curves for CZTI, and at least three matches out of four Veto light curves (Section 4.1). Candidate transients that meet this requirement are flagged as an "event", and entered into an SQL database. Certain basic properties like like number of quadrants and energy bands an event was detected in, their significance, rates above background, time since last SAA, time from next SAA, etc are also calculated and stored in the database. Events having the same trigger time (for instance, if they were detected by two different algorithms) are grouped, and their corresponding event-IDs are stored under a unique trigger-ID in a separate table. Furthermore, the trigger-events which are within 100 seconds of each other are grouped into a "superevent" and assigned a super-ID. These superevents are the final transient candidates, ready for human inspection.

A separate program for plotting is run in parallel which takes input a list of Obs-IDs and fetches all the superevents in those Obs-IDs from the SQL database. For each superevent, it plots detailed time energy histograms, light curves and calculates T_{90} for each temporal binning.

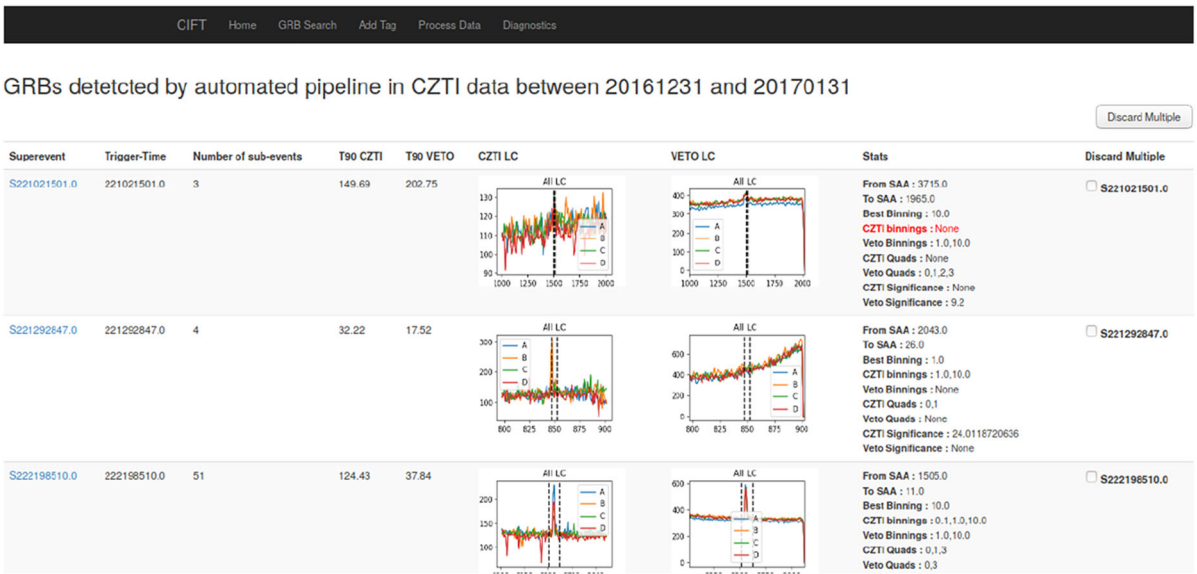
The CZTI Interface For Transients (CIFT) is a Flask[4]-based interface with SQL database as back-

[3] We note that the subsequent arguments become stronger if the actual detected rate is lower as was expected. After completing the search, indeed we found a much lower GRB rate.

[4] https://pypi.org/project/Flask/.



(a) Main page



(b) Scanning page

Figure 3. Screenshots of CIFT, showing functionality of various pages. Panel (**a**): The screenshot of the home page of CIFT, where the human scanner can input dates and the corresponding candidate tag which the scanner wants to see, refer Section 4.5. This main page also allows the user to navigate to other functionalities of the interface where one can add new tags, process the available unprocessed directories and access the diagnostics page, with few clicks. Panel (**b**): The SQL database displays all the candidates of the specified tag from all Obs-IDs contained within the date range specified on the CIFT main page. Each candidate has a dedicated row where the Superevent-ID, trigger time, T_{90} in both CZTI and Veto, number of sub-events are displayed along with relevant statistics like background rate, peak rate, number of quadrants where the candidate was detected for quick reference of the scanner. Each row also has a appropriate light curve thumbnail for both CZTI and Veto for visual inspection, allowing the scanner to discard the very obvious bogus candidates from this page itself (with the help of discard multiple option). This complete list of candidates is sorted in the ascending order of the Superevent-ID. Panel (**c**): Each candidate is linked to their inspection page which displays the break-down of all the computed characteristics shown on the Scanning page. The inspection page also contains links to five different lightcurves for different binnings of CZTI and Veto detectors. Based on the inspection of all these parameters and lightcurves, the scanner can classify the candidate and tag the candidate with the appropriate tag.

(c) Inspection page

Figure 3. *Continued.*

end, available to view the candidate transients (Fig. 3). The interface allows a human scanner to search all superevents by Obs-ID which are displayed in a table on a 'scanning' page (Fig. 3(b)). The scanning page has columns for Superevent-ID, trigger time, T_{90} in CZTI and Veto, number of sub-events, a column displaying relevant statistics like background rate, peak rate, number of quadrants the candidate was detected in, etc. and a check-box option to discard multiple superevents at once if bogus. Each superevent-ID is linked to an inspect page (Fig. 3(c)) which lists all characteristics of the superevent, and of each sub-events contained within it, along with several lightcurves of different binning sizes for CZTI and Veto. After inspection, a human scanner can tag the event with custom tags, including "known", "unknown", "ambiguous", "SAA Tentacle", etc. Superevents can be searched and filtered by tags from the main page (Fig. 3(a)). The CIFT interface also has other features like undertaking triggered searches and a front-end for initiating data processing.

5. Results

We used our framework to search for GRBs in data from 06 October 2015 when CZTI was first powered on, till 10 October 2020 — spanning just over five years of data. "Slew" Obs-IDs are relatively short data sets acquired when AstroSat is slewing from one source to another. These have been excluded from our search. We detected a total of 347 transients in CZTI data by using CIFT. Of these, 41 are GRBs or triggers previously reported by other missions but missed by POC triggered searches or ML pipeline (Section 5.2), while 46 are new discoveries (Section 5.3). In the same five-year span, triggered searches and the ML pipeline have detected 325 GRBs, of which our searches recovered 260. Two of these missed GRBs were in slew orbits. The reasons for missing $\sim$20% GRBs are discussed in Section 5.5.

5.1 *Performance*

The processing code takes less than an hour to search for transient candidates in one month of data (approximately 130 GB). Creating diagnostic plots is a slower process which is spawned in parallel, and takes 3–4 hours to complete. Users remotely connect to the http-based interface for scanning the processed data. Visual examination of candidates from a month of data takes a few hours for an experienced user.

Figure 8 shows the break-up of transient detections by the various algorithms. We see that most transients are detected by all three algorithms, followed by detections in both CS and TN. The TN method is solely responsible for the detection of 15% of Veto transients.

Table 2. Comparison of the three search algorithms running on CZTI and Veto for different classifications of the candidates identified. The 'Candidates' column contains all potential transient candidates identified by our pipeline. The 'Common with triggered or ML searches' column contains all GRBs that were originally detected by triggered or ML searches on CZTI data. The 'Known transients' column contains all transients that had previously been reported by other instruments but had not been identified in CZTI or Veto data. The column 'Discoveries' comprises of all transients that have not been reported by any instrument before. Common events in various methods are shown in Fig. 8.

Algorithm	Candidates	Common with triggered or ML searches	Known transients	New discoveries
CZTI				
Cutoff rate	1290	206	16	29
NSigma	2082	164	7	19
TopN	4199	210	19	30
Veto				
Cutoff rate	10375	191	24	29
NSigma	10625	178	18	23
TopN	13993	222	34	32
Total	22564	260	41	46

Table 2 summarises the performance of all algorithms. We see that there are a large number of false positives, particularly from the Veto detectors. This underscores the need for human vetting of the candidate superevents.

On an average, CIFT flags about 339 candidates per month, adding up to 19628 candidates in 58 months of data. For the months of April and May 2020, we lowered the thresholds to search for even faint bursts associated with the outburst of the galactic magnetar/FRB candidate SGR 1935+2154 (Mereghetti *et al.* 2020). We selected the top 5 peaks in the TN method, and required a coincidence of just 2 bands out of 12 in CR and NS methods. These reduced thresholds increased the number of candidates by a factor of 4.3, giving 2936 candidates in just 2 months.

The most common type of false positives comprised of coincident detections in two Veto quadrants in just a single second, with no discernible signal in adjacent bins. These are most likely particle events, and are rejected. A closely associated class of Veto false positives are events that have a very sharp rise and an exponential decay: again a profile common for particle events. On the other hand, Veto light curves of GRBs that are also detected in CZT detectors show a wider variety. Hence we decided to keep the coincidence threshold for Veto as 3 out of 4 quadrants at the expense of missing possible real transients, and this was the number discussed at the start of Section 4.

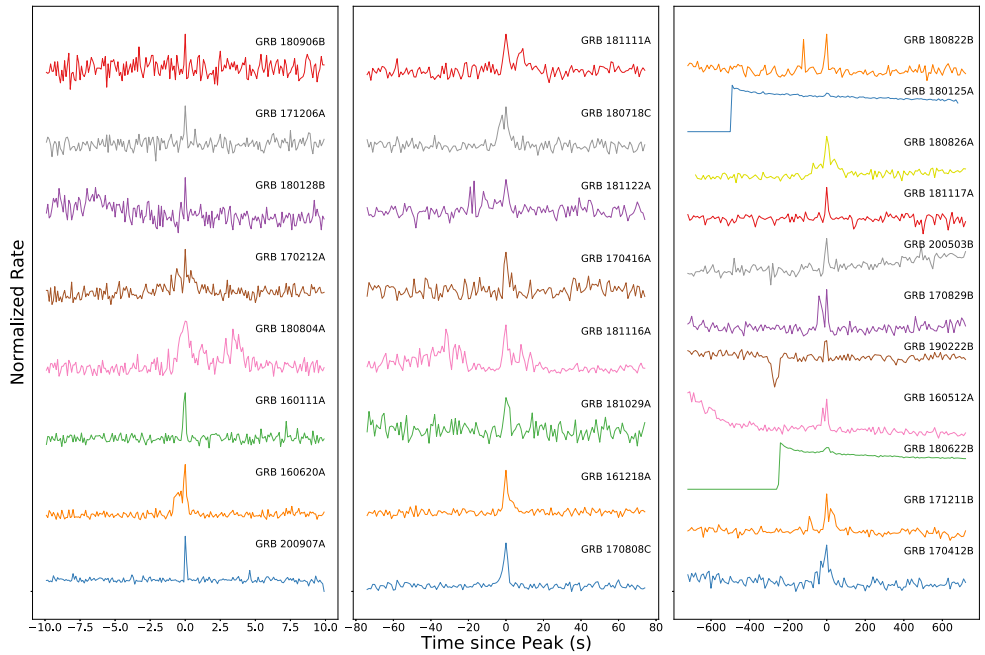
Other large number of bogus detections include false peaks near SAA due to bad de-trending or inadequate SAA masking which can be ruled out during human vetting. In CZTI data, many false events are caused by a single pixel, generating noise events at all energies. Visual examination of the distribution of counts in the detector helps to quickly dismiss these as false positives. If the light curves are well-behaved with no real transients or noise spikes, then the TN algorithm often generates false positives by identifying "outliers" that are completely consistent with background.

As human scanners gain more experience with the pathologies of false positives, we are working to improve automatic rejection of such candidates.

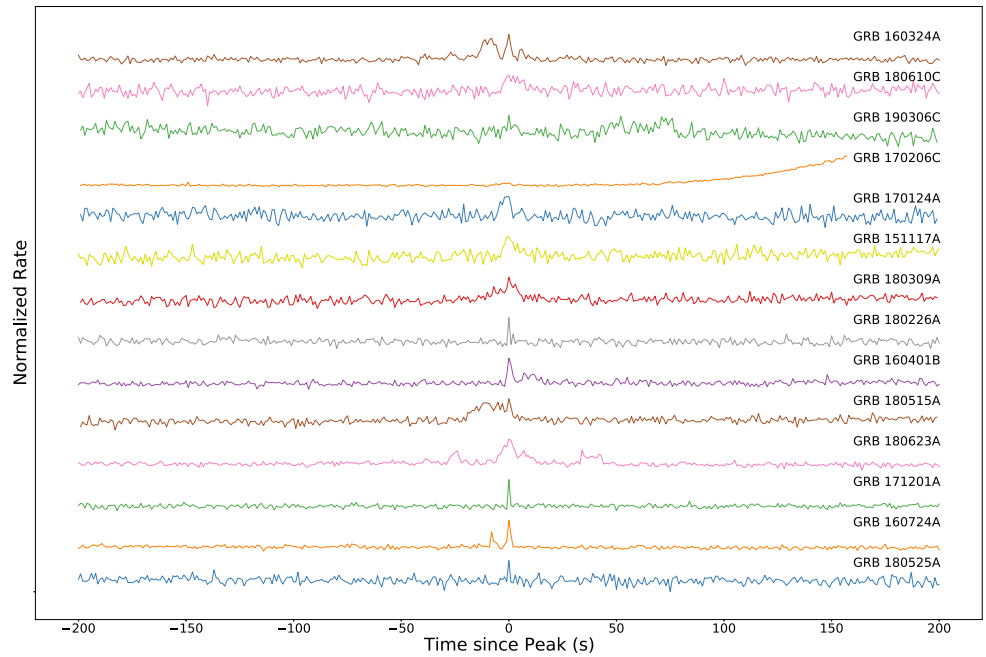
5.2 *Known transients*

We detected 41 transients (referred as 'Known') that had previously been reported by other instruments but had not been identified in CZTI or Veto data (Fig. 4). These transients were matched to earlier reports in GCN Circulars,[5] *Fermi* GBM Burst Catalog[6] and the *Fermi* sub-threshold trigger lists.[7,8] Table 3 lists the key properties of these transients: a superevent ID, standard GRB name, trigger times (UTC), algorithms that detected the transient in CZTI or Veto data,

[5] https://gcn.gsfc.nasa.gov/gcn3_archive.html.
[6] https://heasarc.gsfc.nasa.gov/W3Browse/fermi/fermigbrst.html.
[7] https://gcn.gsfc.nasa.gov/fermi_gbm_subthresh_archive.html.
[8] https://gammaray.nsstc.nasa.gov/gbm/science/sgrb_search.html.



(a) CZTI lightcurves



(b) Veto lightcurves

Figure 4. The normalised lightcurves of GRBs detected by CIFT, that were reported by other instruments but had not been identified in CZTI or Veto data (Section 5.2). Each GRB lightcurve is normalised and labeled with the GRB name. Panel (**a**) shows normalised lightcurves for the GRBs detected in CZTI. The three sub-panels are with 0.1 s, 1 s and 10 s binning respectively, and each sub-panel is ordered by peak count rate above background, increasing from top to bottom. Panel (**b**) shows the normalised lightcurves of GRBs that were detected in Veto. These are plotted with a 1 s binning, and are also ordered by peak count rate above background, increasing from top to bottom.

temporal binning used in analysis, and the peak time (AstroSat time, measured as seconds since UT 2010-01-01, 00:00:00). We then list the calculated parameters: the duration (T_{90}), peak count rates above

background, background count rates and total counts across all quadrants. We prefer using CZTI data to calculate these parameters. Even when our algorithms find a transient only in Veto detectors, we manually check if CZTI data can be used for calculation for uniformity. We use Veto data to calculate transient properties only if the transient is unseen in CZTI light curves. These cases are demarcated clearly in Table 3.

5.3 *CIFT discoveries*

We discovered 46 new transients that have not been reported by any instrument before. As in Section 5.2, we show their light curves in Fig. 5 and list properties in Table 4. Six of these transients have been published already: GRB 180112B (Sharma *et al.* 2018), GRB 190628B (Marathe *et al.* 2019), GRB 191102A (Shenoy *et al.* 2019a), GRB 191105B (Shenoy *et al.* 2019b), GRB 191119A (Shenoy *et al.* 2019c), and GRB 200817B (Shenoy *et al.* 2020).

5.4 *Properties of new transients*

The new transients detected by CIFT (Sections 5.2 and 5.3) span a wide range of properties. The shortest transient was GRB 200907A ($T_{90} = 0.13$ s), while the longest was GRB 180809C with T_{90} of 290 sec. GRB 200510B had the highest count rate above background (6461.6 count/s), while GRB 200906B had the lowest (53.7 count/s). Figure 6 shows the distributions of T_{90}, peak count rate, and total counts for the four classes of transients:

(a) Those reported in the past by CZTI POC,
(b) transients reported by POC which were also found by CIFT,
(c) CIFT-detected transients reported by other instruments, and
(d) new CIFT discoveries.

We observe that all four classes have similar distributions of T_{90}. A notable difference is seen in the total counts: transients with higher number of total CZT counts tend to be easily detected in regular triggered and ML searches. Also, GRBs with low peak count rates are more likely to be found in triggered searches undertaken by the POC but missed by CIFT. Note that although the three classes "POC-GRBs", "Known GRBs", and "New Discovered" are mutually exclusive, the distributions overlap well at the faint end of the distribution.

We find that about 10% of all GRBs detected by CZTI are short GRBs, and the fraction remains the same for the 87 new bursts discovered with CIFT. The fraction of short GRBs is similar to the values for *Swift*-BAT (Lien *et al.* 2016), but smaller than the 26% measured in *Fermi* (von Kienlin *et al.* 2020). Here, we note an important caveat that we draw the line between short and long GRBs at the canonical value of $T_{90} = 2$ s, but it is known that this can be different for different instruments and will have to be measured separately for CZTI. As an illustration, if we adapt the *Fermi* boundary of 6.1 s, we find that about one-third of all CZTI GRBs are short GRBs.

5.5 *Transients missed by CIFT*

Sixty-five GRBs that were found in regular triggered + ML searches were missed in the blind search with CIFT. Two of the missed GRBs were in AstroSat slew orbits which were skipped while processing, as mentioned in Section 5. We analysed the remaining cases to find the reasons why these were missed. The most common reason for the missed GRBs was that the transients were too faint in terms of their peak count rates. For instance, Fig. 7 shows the multi-quadrant, multi-band light curves for GRB 190605A. Visually, it is clear that the GRB is only weakly detected in all three search bands in CZTI data. In order to quantify this further, we calculated the count rates that would have been necessary to flag a data point as an outlier in the peak maps for this orbit. These rates for the CS method are shown with dashed lines, while the 5-σ rates for NS are shown with dotted lines. It is clearly seen that the transient is well below these rates.

Such transients are rather easily confirmed by a human scanner inspecting the spectrogram and finding similar patterns in multiple quadrants. For quantitative analysis with say the CS method, the search window for a triggered search is usually set to 100 s, much smaller than the 4156 s window used in blind searches. This results in a lower cutoff rate, and will make more such fainter transients detectable in the current CIFT framework. Similarly, a smaller search window enables lowering the NS threshold from 5-σ to 4-σ or 3-σ thanks to the fewer data points present, thereby increasing the odds of detecting fainter transients.

Table 3. The table contains the calculated parameters for all 'known GRBs', which are the GRBs that had previously been reported by other instruments but had not been identified in CZTI or Veto data.

SuperID	GRB name	Time (UTC)	Algorithm	Peak time (s)	Bin (s)	T_{90} (s)	Peak rate (cps)	Bkg rate (cps)	Total counts (counts)
Detected: CZTI Analyzed: CZTI		4d							
S196531116.0	GRB160324A[5]	15:58:34	C: CR, TN V: CR, NS, TN	196531117.5	1	48^{+38}_{-17}	120^{+30}_{-14}	$209.2^{+0.9}_{-1.0}$	1889^{+570}_{-589}
S200724370.0	GRB160512A[3†]	04:46:08	C: CR, TN V: CR	200724382.5	1	26^{+6}_{-4}	237^{+43}_{-26}	336^{+2}_{-3}	1854^{+347}_{-382}
S204095177.0	GRB160620A	05:06:15	C: CR, NS, TN V: CR, NS, TN	204095178.4	0.1	$0.99^{+0.62}_{-0.17}$	1385^{+218}_{-222}	447^{+8}_{-10}	714^{+101}_{-103}
S219728856.0	GRB161218A	3:47:34	C: CR, NS, TN V: NS, TN	219728856.1	0.1	6^{+1}_{-1}	457^{+190}_{-18}	423^{+10}_{-12}	1134^{+227}_{-276}
S224556541.0	GRB170212A[2†]	00:48:59	C: CR, NS, TN V: CR, TN	224556540.2	0.1	4^{+2}_{-1}	353^{+150}_{-39}	316^{+8}_{-6}	536^{+135}_{-189}
S229730480.0	GRB170412B[†]	22:01:18	C: CR, NS, TN V: None	229730479.5	1	51^{+7}_{-17}	318^{+44}_{-47}	425^{+3}_{-3}	2589^{+667}_{-689}
S230047201.0	GRB170416A[4†]	13:59:59	C: TN V: CR, NS, TN	230047201.2	0.1	$5.9^{+0.2}_{-1.4}$	288^{+141}_{-26}	333^{+8}_{-7}	551^{+115}_{-135}
S239852052.0	GRB170808C[†]	01:34:10	C: CR, NS, TN V: None	239852052.4	0.1	$4.5^{+0.4}_{-0.5}$	772^{+196}_{-86}	422^{+8}_{-12}	1106^{+185}_{-179}
S250719370.0	GRB171211B[†]	20:16:08	C: CR, TN V: TN	250719471.5	1	135^{+8}_{-8}	278^{+41}_{-37}	358^{+2}_{-2}	4827^{+697}_{-641}
S254815405.0	GRB180128B[1†]	06:03:23	C: CR, TN V: NS, TN	254815407.2	0.1	$5.7^{+0.7}_{-0.8}$	297^{+195}_{-17}	466^{+9}_{-13}	814^{+190}_{-200}
S269585348.0	GRB180718C[†]	04:49:06	C: CR, TN V: None	269585341.5	1	10^{+6}_{-4}	217^{+38}_{-40}	363^{+3}_{-2}	1024^{+157}_{-230}
S271117239.0	GRB180804A[†]	22:20:37	C: CR, NS, TN V: CR, NS, TN	271117238.8	0.1	$5.2^{+0.6}_{-0.8}$	695^{+209}_{-43}	434^{+7}_{-8}	1313^{+152}_{-142}
S272640643.0	GRB180822B[2†]	13:30:41	C: CR, TN V: None	272640638.0	10	126^{+64}_{-110}	72^{+11}_{-10}	$352.4^{+0.4}_{-0.7}$	1589^{+712}_{-705}
S272942427.0	GRB180826A[2†]	01:20:25	C: CR, TN V: CR, NS, TN	272942423.0	10	130^{+11}_{-14}	119^{+12}_{-13}	$493.5^{+0.8}_{-1.1}$	6535^{+565}_{-551}
S279646944.0	GRB181111A[†]	15:42:22	C: TN V: None	279646935.5	1	14^{+3}_{-3}	192^{+41}_{-42}	480^{+3}_{-3}	1138^{+176}_{-210}
S280064235.0	GRB181116A[†]	11:37:13	C: CR, TN V: CR, NS, TN	280064267.5	1	74^{+9}_{-4}	316^{+46}_{-35}	454^{+2}_{-2}	5805^{+590}_{-626}
S280166126.0	GRB181117A[3†]	15:55:24	C: CR, TN V: CR, NS, TN	280166128.5	1	12^{+3}_{-3}	133^{+37}_{-23}	328^{+2}_{-3}	921^{+208}_{-174}
S326222331.0	GRB200503B	17:18:50	C: TN V: None	326222323.5	1	62^{+11}_{-12}	155^{+43}_{-25}	155^{+43}_{-25}	2372^{+656}_{-696}
S337184665.0	GRB200907A[3]	14:24:24	C: CR, NS, TN V: CR, NS, TN	337184664.7	.01	$0.13^{+0.01}_{-0.02}$	4718^{+1176}_{-1079}	476^{+16}_{-47}	260^{+33}_{-29}
Detected: Veto Analyzed: CZTI									
S185452630.0	GRB151117A[†]	10:37:08	C: None V: CR, TN	185452622.8	0.1	5^{+2}_{-1}	267^{+168}_{-9}	383^{+5}_{-8}	493^{+150}_{-155}
S190192926.0	GRB160111A	07:22:04	C: None V: TN	190192925.5	.01		903^{+186}_{-170}	431^{+8}_{-10}	126^{+84}_{-88}
S222932540.0	GRB170124B[4†]	05:42:18	C: None V: CR, TN	222932533.6	0.1	$10.3^{+0.5}_{-1.2}$	264^{+160}_{-17}	328^{+6}_{-6}	897^{+187}_{-154}
S241696600.0	GRB170829B[2†]	09:56:38	C: None V: TN	241696628.5	1	45^{+4}_{-3}	182^{+35}_{-36}	352^{+3}_{-2}	2320^{+297}_{-430}
S250224944.0	GRB171206A[1†]	2:55:42	C: None V: TN	250224944.7	0.1	3^{+1}_{-1}	272^{+170}_{-13}	477^{+9}_{-11}	297^{+156}_{-155}
S254611360.0	GRB180125A[3†]	21:22:38	C: None V: TN	254611351.5	1	24^{+14}_{-6}	$95.5^{+51.8}_{-0.6}$	447^{+3}_{-3}	1352^{+385}_{-394}
S257333099.0	GRB180226A	09:24:57	C: None V: NS, TN	257333099.3	0.1	$0.78^{+0.57}_{-0.53}$	346^{+159}_{-50}	340^{+8}_{-7}	152^{+61}_{-57}
S264915479.0	GRB180525A[†]	03:37:57	C: None V: TN	264915479.1	.01	$0.16^{+0.02}_{-0.05}$	1128^{+798}_{-120}	487^{+20}_{-41}	88^{+27}_{-25}

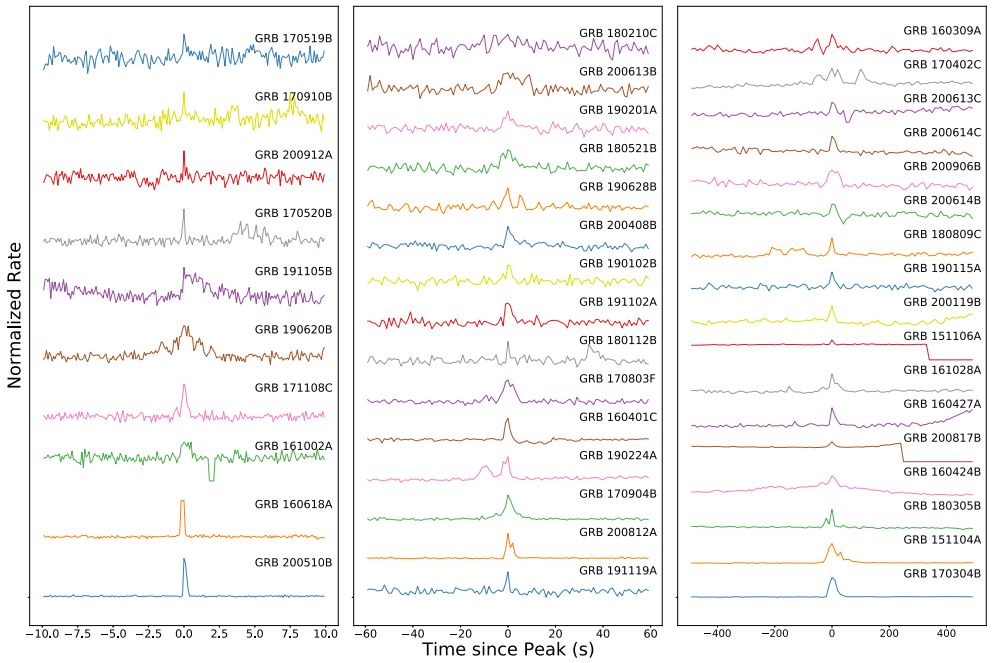
Table 3. *Continued.*

SuperID	GRB name	Time (UTC)	Algorithm	Peak time (s)	Bin (s)	T_{90} (s)	Peak rate (cps)	Bkg rate (cps)	Total counts (counts)
S266353140.0	GRB180610C†	18:58:58	C: None V: CR, TN	266353130.5	1	18^{+1}_{-6}	122^{+47}_{-6}	391^{+3}_{-2}	1164^{+180}_{-185}
S267371600.0	GRB180622B[2]†	13:53:18	C: None V: TN	267371591.5	1	27^{+5}_{-5}	242^{+50}_{-38}	707^{+3}_{-4}	2498^{+457}_{-474}
S273953547.0	GRB180906B†	18:12:25	C: None V: CR, TN	273953548.1	0.1	$5.8^{+0.5}_{-1.9}$	250^{+208}_{-35}	517^{+10}_{-12}	569^{+183}_{-177}
S288536007.0	GRB190222B†	12:53:25	C: None V: CR, NS, TN	288536007.5	1	$10.4^{+0.6}_{-2.9}$	214^{+40}_{-24}	341^{+2}_{-3}	1030^{+136}_{-145}
Detected:	**Veto**								
Analyzed:	**Veto**								
S197170443.0	GRB160401B†	01:34:01	C: None V: CR, NS, TN	197170442.4	1	14^{+14}_{-4}	385^{+60}_{-59}	952^{+5}_{-4}	1626^{+459}_{-581}
S207052803.0	GRB160724A†	10:40:03	C: None V: CR, NS, TN	207052810.8	1	16^{+9}_{-7}	980^{+80}_{-86}	1560^{+5}_{-6}	3098^{+632}_{-666}
S224077190.0	GRB170206C	11:39:48	C: None V: CR, NS, TN	224077218.4	1		218^{+68}_{-30}	1511^{+5}_{-6}	2193^{+594}_{-638}
S249867183.0	GRB171201A	23:33:01	C: None V: TN	249867182.8	1		798^{+76}_{-86}	1555^{+5}_{-7}	853^{+166}_{-185}
S258277398.0	GRB180309A[1]†	07:43:16	C: None V: CR, TN	258277397.7	1	25^{+5}_{-8}	310^{+60}_{-61}	1041^{+3}_{-4}	2484^{+389}_{-416}
S264108750.0	GRB180515A†	19:32:28	C: None V: CR, NS, TN	264108761.5	1	20^{+4}_{-2}	441^{+71}_{-50}	1555^{+5}_{-6}	6091^{+528}_{-582}
S267468139.0	GRB180623A	16:42:17	C: None V: CR, NS, TN	267468140.9	1	64^{+2}_{-4}	518^{+70}_{-48}	1163^{+4}_{-4}	6986^{+773}_{-1088}
S278544286.0	GRB181029A†	21:24:44	C: None V: CR, NS, TN	278544285.6	1	10^{+1}_{-5}	392^{+66}_{-70}	1309^{+5}_{-5}	1286^{+278}_{-307}
S280573750.0	GRB181122A[2]†	09:09:08	C: None V: CR, TN	280573745.5	1	44^{+3}_{-13}	237^{+60}_{-45}	1113^{+4}_{-5}	2634^{+566}_{-623}
S289607820.0	GRB190306C†	22:36:58	C: None V: CR, TN	289607814.3	1		199^{+89}_{-2}	1563^{+6}_{-7}	3597^{+1454}_{-1499}

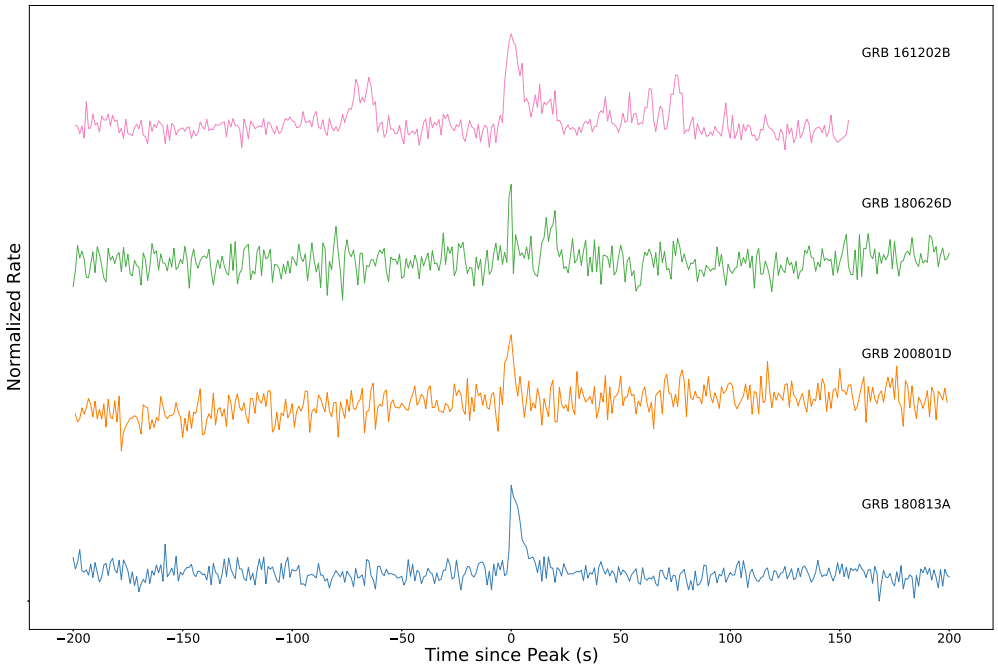
The table is divided into three parts classified by what detector was used to detect ('Detected') and compute the parameters ('Analysed') given in the table. The column 'SuperID' gives the name of the superevent identified by the pipeline. The column 'GRB Name' contains the published name of the GRB, linked to the GRB report (more details in 5.2). Several of these entries in the 'GRB Name' column have a mark against their names, which gives the information on which quadrants were used for calculating the other parameters for that GRB. If there is no mark, then all four quadrants are used. Otherwise, marks '1', '2', '3', '4', and '5' refer to the quadrant sets – 'A, B, C', 'A, B, D', 'A, C, D', 'B, C, D' and 'C, D' respectively. The column 'Algorithm' tells us what algorithms detected the GRB, where 'TN', 'NS', 'CR' stands for the three algorithms – TopN, N-sigma, and Cut-off rate respectively while 'C' and 'V' are the two detectors – CZTI and Veto. The time in *AstroSat* seconds where the GRB was brightest is given in the column 'Peak time'. The bin size, that was used to generate the parameters – T_{90}, Peak count rate above background, Background rate and Total counts, is mentioned in the column 'Bin'. GRBs marked with a dagger (†) were obtained from the FERMIGBRST Catalogue (von Kienlin *et al.* 2020; Gruber *et al.* 2014; von Kienlin *et al.* 2014; Narayana Bhat *et al.* 2016).

6. Conclusions and future work

CZTI has proven itself to be a sensitive transient detector, but our searches had largely been limited to triggered searches. The ML pipeline (Abraham *et al.* 2019) was the first major step towards detection of new transients with CZTI. The development of these algorithms, software, and the CIFT interface provide us with a powerful tool to extend our work further. Here, we have demonstrated the utility of this tool with the discovery of 87 new transients that had been missed by previous searches, including 46 transients that had not been detected by any mission to date. This brings the total CZTI tally to 412 GRBs in the first five years of its operation since launch, or about ~ 83 per year. For comparison, *Swift* BAT detects ~ 92 GRBs per year from on-board triggers (Lien *et al.* 2016), while *Fermi* GBM detects ~ 235 GRBs per year (von Kienlin *et al.* 2020).



(a) CZTI lightcurves



(b) Veto lightcurves

Figure 5. The normalised lightcurves of GRBs discovered by CIFT that have not been reported by any instrument before (Section 5.3). Each GRB lightcurve is normalised and labeled with the GRB name. Panel (**a**) shows normalised lightcurves for the GRBs detected in CZTI. The three sub-panels are with 0.1 s, 1 s and 10 s binning respectively, and each sub-panel is ordered by peak count rate above background, increasing from top to bottom. Panel (**b**) shows the normalised light curves of GRBs that were detected in Veto. These are plotted with a 1 s binning, and are also ordered by peak count rate above background, increasing from top to bottom.

The CIFT framework is constantly evolving. It has been designed to make it easy to incorporate new features including search algorithms. We are working on metrics to quantify the statistical significance of a

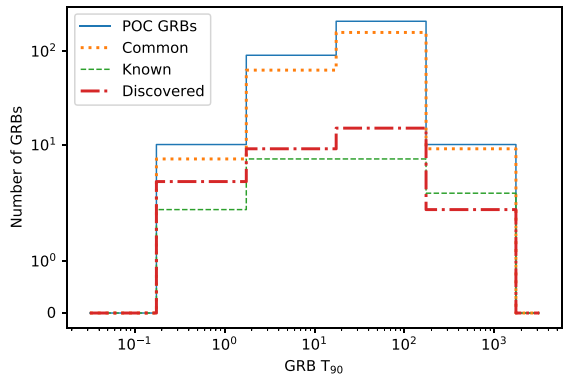
Table 4. The table contains the calculated parameters for all 'Discovered GRBs', which are all GRBs that have not been reported by any instrument before.

SuperID	GRB name	Time (UTC)	Algorithm	Peak time (s)	Bin (s)	T_{90} (s)	Peak rate (cps)	Bkg rate (cps)	Total counts (counts)
Detected:	**CZTI**								
Analyzed:	**CZTI**								
S184303433.0	GRB151104A	3:23:51	C: CR, NS, TN V: CR, NS, TN	184303438.5	1	68^{+2}_{-4}	911^{+62}_{-43}	474^{+3}_{-4}	28049^{+744}_{-827}
S184512194.0	GRB151106A[3]	13:23:12	C: CR, NS, TN V: None	184512190.5	1	$39.4^{+0.8}_{-25.5}$	203^{+39}_{-25}	319^{+2}_{-3}	1596^{+285}_{-292}
S195263948.0	GRB160309A[5]	23:59:06	C: CR, TN V: None	195263943.0	10	67^{+27}_{-16}	62^{+8}_{-9}	$211.3^{+0.4}_{-0.6}$	2034^{+521}_{-618}
S197184964.0	GRB160401C[1]	05:36:02	C: CR, NS, TN V: CR, NS, TN	197184965.2	0.1	$3.1^{+0.9}_{-1.0}$	588^{+189}_{-20}	386^{+10}_{-10}	1090^{+152}_{-180}
S199212123.0	GRB160424B	16:42:01	C: CR, NS, TN V: None	199212122.5	1	$43.6^{+0.7}_{-2.2}$	691^{+54}_{-62}	554^{+4}_{-3}	4093^{+342}_{-511}
S199449121.0	GRB160427A	10:31:59	C: CR, TN V: CR, NS, TN	199449124.5	1	34^{+3}_{-12}	375^{+46}_{-50}	476^{+4}_{-3}	2669^{+317}_{-372}
S203963626.0	GRB160618A	16:33:44	C: CR, NS, TN V: CR, NS, TN	203963626.6	.01	$0.27^{+0.01}_{-0.01}$	4918^{+1328}_{-366}	444^{+17}_{-48}	877^{+61}_{-61}
S215358438.0	GRB161028A	13:47:16	C: CR, TN V: None	215358429.5	1	$25.7^{+0.6}_{-0.3}$	284^{+44}_{-36}	434^{+3}_{-4}	1883^{+267}_{-271}
S218385140.0	GRB161202B[4]	14:32:27	C: CR, TN V: CR, NS, TN	218385206.5	1	146^{+1}_{-4}	131^{+35}_{-20}	326^{+2}_{-2}	3346^{+767}_{-848}
S226302000.0	GRB170304B	05:39:58	C: CR, NS, TN V: CR, NS, TN	226302047.5	1	41^{+2}_{-3}	5091^{+120}_{-126}	486^{+2}_{-2}	23464^{+495}_{-511}
S232990228.0	GRB170520B	15:30:26	C: CR V: CR, NS, TN	232990228.1	0.1	$2.3^{+0.8}_{-0.5}$	506^{+176}_{-61}	490^{+11}_{-10}	551^{+101}_{-162}
S239432831.0	GRB170803F	05:07:09	C: CR, NS, TN V: CR, NS, TN	239432833.4	0.1	$6.0^{+0.3}_{-0.4}$	344^{+172}_{-68}	452^{+8}_{-13}	1494^{+209}_{-227}
S242255598.0	GRB170904B	21:13:16	C: CR, NS, TN V: NS, TN	242255598.6	0.1	$3.4^{+0.4}_{-0.5}$	966^{+208}_{-175}	553^{+10}_{-12}	1424^{+199}_{-160}
S247846852.0	GRB171108C[3]	14:20:50	C: CR, NS, TN V: None	247846851.7	0.1	$0.88^{+0.48}_{-0.26}$	915^{+183}_{-143}	374^{+6}_{-9}	326^{+69}_{-63}
S253483949.0	GRB180112B	20:12:27	C: TN V: CR, TN	253483948.3	0.1	9^{+2}_{-3}	335^{+178}_{-3}	409^{+5}_{-6}	794^{+210}_{-159}
S257929386.0	GRB180305B	07:03:04	C: CR, NS, TN V: CR, NS, TN	257929428.5	1	32^{+2}_{-5}	769^{+56}_{-59}	436^{+2}_{-3}	5170^{+374}_{-315}
S264583245.0	GRB180521B	07:20:43	C: CR, TN V: None	264583239.5	1	15^{+4}_{-9}	167^{+46}_{-24}	529^{+3}_{-3}	1026^{+287}_{-289}
S271507716.0	GRB180809C	10:48:34	C: CR, NS, TN V: CR, TN	271507720.0	10	290^{+22}_{-52}	132^{+13}_{-14}	535^{+1}_{-2}	8438^{+1314}_{-1464}
S286741010.0	GRB190201A[2]	18:16:48	C: CR, TN V: CR, TN	286741004.5	1	$8.0^{+0.4}_{-2.9}$	136^{+38}_{-34}	369^{+2}_{-4}	558^{+122}_{-119}
S288677154.0	GRB190224A	04:05:52	C: CR, NS, TN V: None	288677164.5	1	$12.9^{+0.4}_{-0.4}$	843^{+63}_{-67}	606^{+4}_{-5}	5397^{+254}_{-261}
S298744222.0	GRB190620B	16:30:20	C: CR, NS, TN V: None	298744221.9	0.1	4^{+2}_{-1}	824^{+238}_{-55}	549^{+12}_{-15}	1504^{+362}_{-426}
S299391822.0	GRB190628B[3]	04:23:40	C: CR, TN V: None	299391815.5	1.0	32^{+15}_{-24}	196^{+40}_{-42}	405^{+3}_{-3}	1177^{+410}_{-513}
S310393113.0	GRB191102A[1]	12:18:31	C: CR, TN V: None	310393104.5	0.1	$4.2^{+0.8}_{-1.2}$	279^{+189}_{-25}	392^{+7}_{-10}	583^{+126}_{-124}
S310630593.0	GRB191105B	06:16:31	C: CR, NS, TN V: CR, NS, TN	310630601.8	0.1	$13.2^{+0.4}_{-0.6}$	665^{+202}_{-77}	739^{+7}_{-12}	2200^{+340}_{-343}
S311856067.0	GRB191119A	10:41:05	C: CR, NS, TN V: CR, NS, TN	311856067.0	.01	$0.15^{+0.03}_{-0.02}$	2702^{+1130}_{-230}	449^{+18}_{-42}	270^{+39}_{-34}
S317162956.0	GRB200119B	20:49:13	C: CR, TN V: CR, NS, TN	317162955.5	1		195^{+52}_{-22}	524^{+4}_{-4}	1593^{+1232}_{-1522}
S324009902.0	GRB200408B	2:44:59	C: CR, NS, TN V: None	324009901.5	1		217^{+45}_{-43}	494^{+4}_{-4}	590^{+731}_{-531}

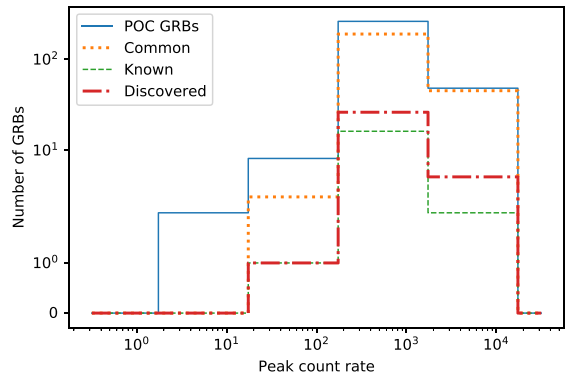
Table 4. *Continued.*

SuperID	GRB name	Time (UTC)	Algorithm	Peak time (s)	Bin (s)	T_{90} (s)	Peak rate (cps)	Bkg rate (cps)	Total counts (counts)
S326787080.0	GRB200510B	6:11:17	C: CR, NS, TN V: None	326787121.8	.01	$0.29^{+0.01}_{-0.01}$	8318^{+1896}_{-366}	520^{+20}_{-50}	1695^{+84}_{-75}
S329736162.0	GRB200613B	09:22:40	C: CR, TN V: CR, NS, TN	329736159.5	1	11^{+2}_{-1}	132^{+58}_{-3}	618^{+4}_{-3}	1190^{+208}_{-236}
S329806842.0	GRB200614C[2]	05:00:41	C: CR, TN V: CR, NS, TN	329806843.5	1	24^{+7}_{-10}	104^{+44}_{-8}	433^{+3}_{-3}	1257^{+288}_{-306}
S334929280.0	GRB200812A	11:54:37	C: CR, NS, TN V: CR, NS, TN	334929324.5	1	10^{+2}_{-5}	1677^{+72}_{-81}	449^{+5}_{-3}	4415^{+194}_{-224}
S335340170.0	GRB200817B	06:02:48	C: CR, NS, TN V: CR, NS, TN	335340230.5	1.0	23^{+10}_{-5}	455^{+48}_{-43}	490^{+3}_{-2}	3847^{+378}_{-507}
Detected:	**Veto**								
Analyzed:	**CZTI**								
S213125894.0	GRB161002A	17:38:12	C: None V: CR, TN	213125894.5	.01	$0.87^{+0.14}_{-0.31}$	991^{+792}_{-56}	413^{+34}_{-41}	235^{+116}_{-105}
S228861170.0	GRB170402C[1]	20:32:48	C: None V: CR, TN	228861062.0	10	119^{+68}_{-41}	82^{+11}_{-9}	$360.9^{+0.8}_{-0.9}$	5072^{+1150}_{-1232}
S242743550.0	GRB170910B	12:45:48	C: None V: TN	242743547.1	0.1	$2.6^{+0.2}_{-1.8}$	400^{+164}_{-94}	471^{+9}_{-10}	259^{+107}_{-106}
S255979768.0	GRB180210C[3]	17:29:26	C: None V: CR, NS, TN	255979759.5	1	7^{+3}_{-1}	43^{+45}_{-6}	350^{+3}_{-3}	110^{+462}_{-511}
S267689714.0	GRB180626D	06:15:12	C: None V: CR	267689735.5	1	23^{+6}_{-3}	133^{+51}_{-10}	529^{+3}_{-3}	1082^{+314}_{-344}
S284139623.0	GRB190102B[1]	15:40:21	C: None V: TN	284139622.4	0.1	$3.1^{+0.5}_{-0.6}$	252^{+154}_{-12}	345^{+7}_{-10}	309^{+88}_{-86}
S285215230.0	GRB190115A[2]	02:27:08	C: None V: CR, TN	285215220.5	1	19^{+1}_{-2}	162^{+40}_{-24}	384^{+2}_{-3}	1225^{+187}_{-194}
S329776004.0	GRB200613C[3]	20:26:42	C: None V: NS	329776008.5	1	23^{+4}_{-10}	97^{+35}_{-17}	334^{+3}_{-3}	925^{+224}_{-247}
S329801810.0	GRB200614B[4]	03:36:47	C: None V: NS, TN	329801807.5	1	$38.3^{+0.5}_{-24.2}$	112^{+47}_{-6}	396^{+2}_{-3}	1306^{+379}_{-294}
S337054619.0	GRB200906B	02:16:56	C: None V: CR, NS	337054499.5	1	35^{+14}_{-9}	110^{+45}_{-6}	462^{+2}_{-2}	2033^{+690}_{-564}
S337615721.0	GRB200912A	14:08:38	C: None V: CR, NS, TN	337615720.8	0.1	$2.6^{+0.0}_{-1.3}$	415^{+169}_{-56}	497^{+9}_{-10}	396^{+83}_{-112}
Detected:	**Veto**								
Analyzed:	**Veto**								
S232918513.0	GRB170519B	19:35:11	C: None V: CR, TN	232918512.2	1	14^{+6}_{-4}	377^{+71}_{-64}	1583^{+6}_{-7}	2251^{+400}_{-448}
S271825254.0	GRB180813A	03:00:52	C: None V: CR, NS, TN	271825253.3	1	16^{+9}_{-7}	506^{+72}_{-60}	1464^{+5}_{-6}	3144^{+434}_{-493}
S333985240.0	GRB200801D	13:40:38	C: None V: CR, TN	333985233.0	1	6^{+1}_{-3}	283^{+72}_{-51}	1687^{+6}_{-6}	1198^{+251}_{-272}

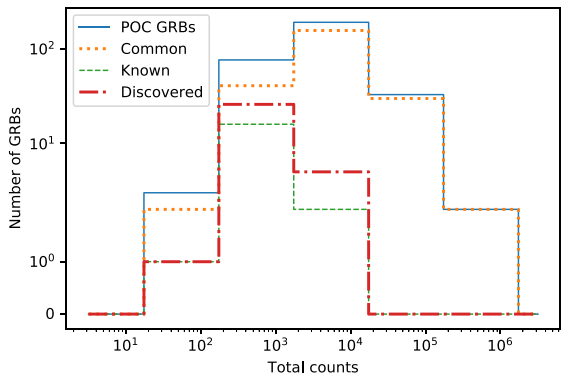
The table is also divided into three parts classified by what detector was used to detect ('Detected') and compute the parameters ('Analysed') given in the table. The column 'SuperID' gives the name of the superevent identified by the pipeline. Several of these entries in the 'GRB Name' column have a mark against their names that tells what quadrants were used for calculating all other parameters for that GRB. If there is no mark, then all four quadrants are used. Otherwise, marks '1', '2', '3', '4', and '5' refer to the quadrant sets – 'A, B, C', 'A, B, D', 'A, C, D', 'B, C, D' and 'C, D' respectively. The column 'Algorithm' tells us what algorithms detected the GRB, where 'TN', 'NS', 'CR' stands for the three algorithms – TopN, N-sigma, and Cut-off rate respectively whereas 'C' and 'V' are the two detectors – CZTI and Veto. The time in AstroSat seconds where the GRB was brightest is given in the column 'Peak time'. The bin size, that was used to generate the parameters – T_{90}, Peak count rate above background, Background rate and Total counts, is mentioned in the column 'Bin'.



(a)



(b)



(c)

Figure 6. (**a**) Distribution of T_{90} values of all CZTI GRBs. (**b**) Distribution of peak count rates of transients in CZT detectors. Note that high peak rates are often obtained for short duration transients analysed with 0.1 s binning. (**c**) Distribution of total counts in CZT detectors. Comparing the distributions of the duration (T_{90}), peak count rate, and total counts in three search classes. Blue lines ("POC GRBs") denote transients detected in regular triggered searches and ML pipeline searches. Orange lines ("Common GRBs") denote the transients CIFT detected among the "POC GRBs". Green lines ("known GRBs") denote transients that have been reported by other instruments (Section 5.2) but missed by POC searches or ML pipeline, while red lines ("Discovered GRBs") denote the new transients discovered with CIFT.

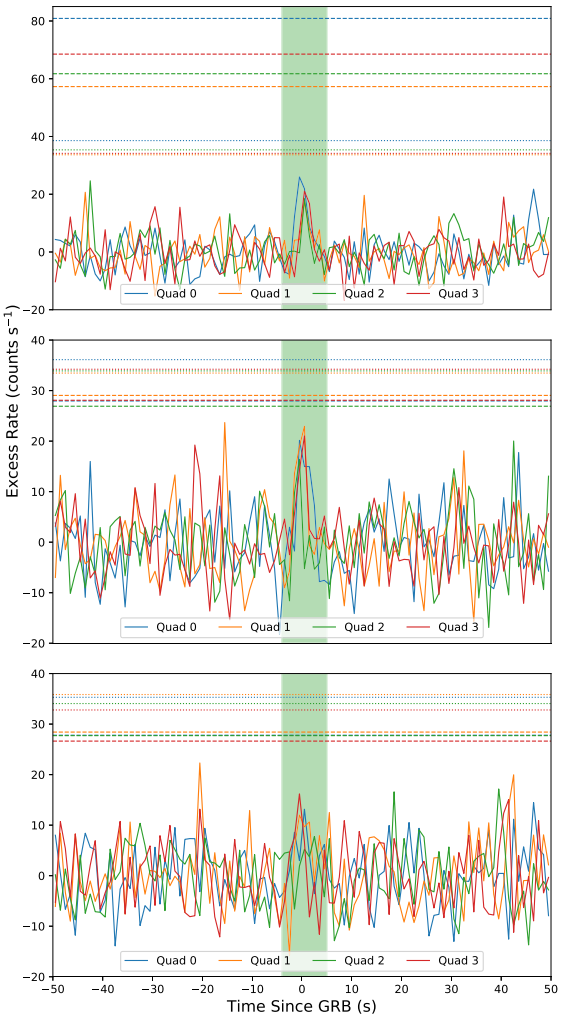
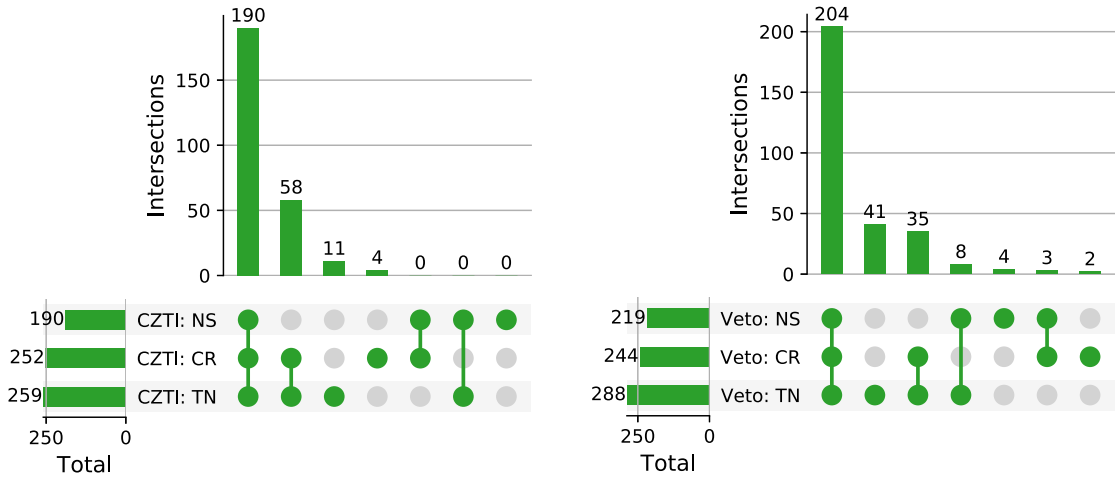


Figure 7. Diagnostic light curves for GRB 190605A. *Top panel*: 20–50 keV light curves for all four CZTI quadrants. The shaded green region denotes the GRB. Dashed and dotted lines denote the outlier threshold for CS and NS methods respectively, for each of the four quadrants. *Middle panel*: Same as top panel, but for 50–100 keV. *Bottom panel*: Same as top panel, but for 100–200 keV. The transient light curve looks similar in all quadrants, but it is too faint to qualify as an outlier in any of the methods.

transient, so that we can lower the FPP. We have developed and tested a new search based on Bayesian Blocks (BB; Scargle *et al.* 2012). We use the `astropy.stats.bayesian_blocks` module to obtain block representations of de-trended light curves, and search for blocks that are 3-σ outliers. These outliers then form the peak maps discussed in Section 4.1, so the BB search can easily be integrated into CIFT as a fourth algorithm. Preliminary testing has shown promising results with significantly lower false positive rates as compared to other algorithms. We will now run the BB search on the full data set.



(a) Statistics of transients detected in CZTI.

(b) Statistics of transients detected in Veto detectors.

Figure 8. Statistics of transients detected in CZTI (Panel (**a**)) and Veto (Panel (**b**)) detectors. The three bottom rows show the number of transients detected by each method: for instance, the N-sigma method detected 115 transients in CZTI. The bar charts at the top show overlaps between various combinations of methods. We see that of the 263 transients detected by CZTI detectors, 190 were detected by all three methods, while another 58 were detected by both the "cutoff rates" and "Top-N" methods. Among the 297 transients found in the Veto detector, 204 transients were detected by all three methods, 41 were detected only by the "Top-N" method, and 35 were detected by both the "cutoff rates" and "Top-N" methods.

Searches for fast transients also stand to benefit from other developments in CZTI data processing. New methods for rejecting noise from raw data (Ratheesh *et al.* 2021) are improving the quality of light curves. These promise to lower the cut-off rates for CS by a factor of a few and will give a proportional boost to the count rate sensitivity of CZTI. Another notable change to be introduced is the non-removal of Veto-tagged events. The default CZTI pipeline attributes coincident events between CZT and Veto detectors to charged particles, and discards them. In case of bright GRBs, large numbers of photons are incident both on CZT and Veto detectors, greatly increasing the chance coincidence rates. Since these are real GRB photons which should not be discarded, future CIFT-based searches will disable Veto-event filtering.

We have also added functionality to undertake specialised searches for X-ray counterparts to Fast Radio Bursts (FRBs) and Gravitational Wave (GW) sources. For instance, the magnetar source SGR 1935+2154 became active in early 2020, creating a series of bursts including one coincident with Fast Radio Burst (Li *et al.* 2020). We used the CIFT interface to efficiently search CZTI data for any bursts from this source. A first blind search was conducted with the default thresholds and we found three bursts, coincident with times reported by other instruments. We then lowered the search thresholds and found an additional four bursts, corresponding to those reported by other missions. We are in the process of analysing properties of these CZTI-detected bursts, and the results will be reported separately (Raman *et al.*, in prep.). We have also incorporated the ability to process GW localisation maps to calculate direction-dependent sensitivity. These features will streamline and boost the effort to search for X-ray counterparts to GW sources from the third observing run of advanced gravitational wave detectors (Abbott *et al.* 2020).

Acknowledgements

CZT-Imager is built by a consortium of Institutes across India. The Tata Institute of Fundamental Research, Mumbai, led the effort with instrument design and development. Vikram Sarabhai Space Centre, Thiruvananthapuram provided the electronic design, assembly and testing. ISRO Satellite Centre (ISAC), Bengaluru provided the mechanical design, quality consultation and project management. The Inter University Centre for Astronomy and Astrophysics (IUCAA), Pune did the Coded Mask design, instrument calibration, and Payload Operation Centre. Space Application Centre (SAC) at Ahmedabad

provided the analysis software. Physical Research Laboratory (PRL) Ahmedabad, provided the polarisation detection algorithm and ground calibration. A vast number of industries participated in the fabrication and the University sector pitched in by participating in the test and evaluation of the payload. The Indian Space Research Organisation funded, managed and facilitated the project. This work utilised various software including Python, AstroPy (Robitaille *et al.* 2013), NumPy (van der Walt *et al.* 2011), Matplotlib (Hunter 2007), https://github.com/jnothman/upsetplot/ UpSetPlot (Lex *et al.* 1983), and ngrok.

References

Abbott R., Abbott T. D., Abraham S. *et al*. 2020, GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run, arXiv:2010.14527

Abraham S., Mukund N., Vibhute A. *et al*. 2019, ArXiv e-prints, 1906.09670, arXiv:1906.09670

Anumarlapudi A., Bhalerao V., Tendulkar S. P., Balasubramanian A. 2020, Astrophys. J., 888, 40

Bhalerao V., Bhattacharya D., Rao A. R., Vadawale S. 2015, GRB Coordinates Network, 18422, 1

Bhalerao V., Kumar V., Bhattacharya D., Rao A. R., Vadawale S. 2016, GRB Coordinates Network, 19519, 1

Bhalerao V., Kasliwal M., Bhattacharya D. *et al*. 2017a, Astrophys. J., 845, arXiv:1706.00024

Bhalerao V., Bhattacharya D., Vibhute A. *et al*. 2017b, J. Astrophys. Astr., 38, 31

Chattopadhyay T., Gupta S., Sharma V. *et al*. 2021, J. Astrophys. Astr., 42, https://doi.org/10.1007/s12036-021-09718-2

Gruber D., Goldstein A., Weller von Ahlefeld V. *et al.* 2014, 211, 12

Gupta S., Sharma V., Bhattacharya D. *et al*. 2020, GRB Coordinates Network, 28451, 1

Hunter J. D. 2007, Computing in Science & Engineering, 9, 90

Lex A., Gehlenborg N., Strobelt H., Vuillemot R., Pfister H. 2014, IEEE Trans. on Visualization and Computer Graphics, 20, 1983

Li C. K., Lin L., Xiong S. L. *et al*. 2020, arXiv e-prints, arXiv:2005.11071

Lien A., Sakamoto T., Barthelmy, S. D. *et al*. 2016, Astrophys. J., 829, 7

Marathe A., Sharma Y., Bhalerao V. *et al*. 2019, GRB Coordinates Network, 24972, 1

Mate S., Bhalerao V., Bhattacharya D. *et al*. 2017, GRB Coordinates Network, 20796, 1

Mate S., Chattopadhyay T., Bhalerao V. *et al*. 2021, submitted to J. Astrophys. Astr., 42, https://doi.org/10.1007/s12036-021-09763-X

Mereghetti S., Savchenko V., Ferrigno C. *et al*. 2020, Astrophys. J. Lett., 898, L29

Narayana Bhat P., Meegan C. A., von Kienlin A. *et al*. 2016, 223, 28

Rao A. R., Chand V., Hingar M. K. *et al*. 2016, Astrophys. J., 833, 86

Ratheesh A., Rao A., Mithun N. *et al*. 2021, J. Astrophys. Astr., 42, https://doi.org/10.1007/s12036-021-09716-4

Robitaille T. P., Tollerud E. J., Greenfield P. *et al*. 2013, A&A, 558, A33

Scargle J. D., Norris J. P., Jackson B., Chiang J. 2012, Astrophys. J., 764, 167

Sharma Y., Bhalerao V., Khanam T. *et al*. 2018, GRB Coordinates Network, 23511, 1

Shenoy V., Aarthy E., Bhalerao V. *et al*. 2020, GRB Coordinates Network, 27315, 1

Shenoy V., Sharma Y., Bhalerao V. *et al*. 2019a, GRB Coordinates Network, 26378, 1

Shenoy V., Sharma Y., Bhalerao V. *et al*. 2019b, GRB Coordinates Network, 26376, 1

Shenoy V., Sharma Y., Bhalerao V. *et al*. 2019c, GRB Coordinates Network, 26268, 1

Shenoy V., Bhalerao, V., Gupta, S. *et al*. 2020, GRB Coordinates Network, 28354, 1

Singh K. P., Tandon S. N., Agrawal P. C. *et al*. 2014, in Takahashi T., den Herder J.-W. A., Bautz M., eds, Space Telescopes and Instrumentation 2014: Ultraviolet to Gamma Ray, vol. 9144, 91441S

van der Walt S., Colbert S. C., Varoquaux G. 2011, Computing in Science & Engineering, 13, 22

von Kienlin A., Meegan C. A., Paciesas W. S. *et al*. 2014, 211, 13

von Kienlin A., Meegan C. A., Paciesas W. S. *et al*. 2020, Astrophys. J., 893, 46