A TRIGGER CLASSIFICATION ALGORITHM FOR THE GLAST BURST MONITOR

David Perrin

Eli Sidman

Marshall Space Flight Center

July 29, 2004

Reviewed by NASA-USRP Mentor
Charles Meegan

Science Directorate
Gamma Ray Astronomy Department

Abstract

Gamma ray bursts (GRBs) are the most powerful events in the universe, occurring in random locations in the sky several times a day. To better understand gamma ray bursts, it is necessary to locate them quickly and accurately because their intensities decrease very rapidly. This will be possible with communication between the Gamma Ray Large Area Space Telescope's (GLAST) two instruments, the Large Area Telescope (LAT), and the GLAST Burst Monitor (GBM). The GBM will detect gamma ray events using its full sky field of view, and an onboard program will calculate the probability that the gamma radiation corresponds to a specific physical event. The possible events are GRBs, solar flares, hard X-ray transients, soft gamma repeaters, magnetospheric particles, and Cygnus X-1. If the probability that the event is a GRB is above a defined confidence level, and is sufficiently intense, the GLAST satellite will quickly turn so the LAT can look at the GRB with better resolution. The parameters used to distinguish events in the Bayesian analysis were the McIlwain magnetic L coordinate, geocenter angle, and the right ascension and declination of the event in comparison to known sources. Probability distributions for these parameters were derived using data from the Burst and Transient Source Experiment. The program was tested with independent data when possible, and the best confidence level for GRBs was determined to be 70%. At this confidence level, the algorithm had a 95% success rate identifying GRBs and an 86% success rate of not identifying other events as GRBs.

1. GAMMA-RAY BURSTS

Gamma ray bursts (GRBs) are the most powerful events in the universe, appearing in random locations throughout the entire sky. GRBs generally fall into two categories, short and long, and are believed to have different causes. Short bursts, lasting under two seconds, are believed to be the result of the merger of two neutron stars. As the two stars orbit each other, they slowly radiate away energy and spiral inwards. When they finally merge, they form a black hole, releasing an enormous amount of energy, mostly out along its axis of rotation in both directions. Long bursts are thought to be caused by a hypernova, also known as a collapsar. As a giant star begins to collapse in on itself, a black hole forms in its interior, while the outside layers' angular momentum temporarily keeps them from falling in. The matter along the axis of rotation has less angular momentum and falls in more easily, collapsing the star into a doughnut shape. Eventually, from a few seconds to over a hundred seconds later, the rest of the star falls victim to its own black hole. The entire process spews vast amounts of energy out along its axis in both directions. The common factor in both cases is the formation of a black hole. It is estimated that about three GRBs are visible from Earth per day. However, because the radiation is beamed into a narrow cone, there could easily be hundreds of GRBs throughout the universe every day. Gamma rays cannot penetrate the earth's atmosphere, and thus the observing of cosmic gamma rays and GRBs must be done from satellites or from high altitude balloons.

From 1991 to 2000, the Burst and Transient Source Experiment (BATSE) aboard the Compton Gamma Ray Observatory (CGRO) observed 2704 gamma ray bursts, as well as many other gamma ray events. BATSE observed in four energy ranges: Channel 1 observed 25-50 keV, Channel 2 observed 50-100 keV, Channel 3 observed 100-300 keV, and Channel 4 observed above 300 keV. BATSE showed that the GRBs came from random locations around the sky, rather than only the galactic plane. This strongly suggested that GRBs are at cosmological distances, rather than galactic. However, there are still many questions to be answered about what causes GRBs, and further study of these events is required.

There are several events other than GRBs that can cause a rise in gamma radiation high enough above the cosmic background to trigger our detectors. Solar flares often produce gamma rays. Some neutron stars, known as magnetars, that produce gamma rays are called soft gamma repeaters (SGRs). Hard X-ray transients are binary neutron star systems that sometimes go through periods of gamma ray outbursts. Magnetospheric particle precipitation events are caused by particles caught in the Van Allen belts that produce gamma rays. There are two types of these particle events. Local particle events occur at the spacecraft and cause a flood of counts in all of the detectors. Distant particle events occur near the Earth's horizon and thus are seen by only a few detectors, much like other distant events. Finally, Cygnus X-1 is a black hole system that occasionally goes through gamma ray outbursts.

A typical GRB lasts only a few seconds, and then leaves a faint X-ray or optical afterglow that decays to invisibility over several days to weeks depending on the intensity of the burst and the sensitivity of the observing equipment. With BATSE, a team of scientists reviewed the triggers each day to determine whether each event was a GRB or not. This meant that a significant delay occurred before ground-based observation could begin searching for optical and radio afterglows. For this reason, fast and accurate locations of GRBs are needed.

The Gamma Ray Large Area Space Telescope (GLAST) is currently being developed to further the knowledge of GRBs and other gamma ray sources. The GLAST will contain two instruments, the GLAST Burst Monitor (GBM) for low energy ranges, and the Large Area Telescope (LAT) for high energy ranges. The GBM will detect gamma rays in the energy range of a few keV up to about 30 MeV. The GBM will consist of 12 Sodium iodide (NaI) scintillation crystals and 2 Bismuth germanate (BGO) scintillation crystals. The NaI crystals will detect gamma rays in the range of a few keV to about one MeV. The BGO crystals will overlap the NaI crystals and the LAT, detecting gamma rays in the range of 150 keV to about 30 MeV (Gehrels, 2002, Instrument Design). The LAT will detect gamma rays in the range of 20 MeV to 300 GeV.

2. EVENT CLASSIFICATION

The GBM will trigger on increases in gamma radiation from any source in the entire sky not blocked by the earth. An onboard program will compute the probability that each possible type of physical event, such as a GRB or solar flare, caused the trigger. If it is highly probable that the event is a GRB, and the event is intense enough, the spacecraft will swing quickly to allow the LAT to observe the GRB with greater resolution and accuracy. The LAT has a much narrower field of view than the GBM, so this turning of the spacecraft is necessary to put the GRB in its field of view. This will allow the GRB to be monitored for much longer, and at more energy ranges than was possible with BATSE-observed bursts. This paper will describe the process of developing the program and its effectiveness with test data.

2.1 Filtered Events

Some properties of the detected events will be used to filter out particle events before the probability calculations. The opposite over maximum detector count ratio, calculated by dividing the maximum trigger count in Channel 2+3 into the trigger count of the detector on the opposite side of the spacecraft, will be used to eliminate local particle events. The eight BATSE detectors were arranged in an octahedron around the satellite, so each detector had an opposite counterpart. This was an important distinguisher of particle events. This ratio for particle events was closer to one than other events because particle events occur close to the spacecraft and send showers of gamma rays all around it (see Figure 1). Events from far away are detected strongly by only one detector, or two that are adjacent. Maximum over minimum detector count ratio was also calculated, but turned out to be a much less reliable parameter than maximum over opposite. Maximum over minimum was the ratio of the highest trigger count in Channel 2+3 to the lowest trigger count in Channel 2+3. The geocenter angle helps distinguish distant particle events. It is known that most distant particle events come from the geocenter angle of the earth's atmosphere, 72 degrees, so events at or below the horizon are also filtered out.

2.2 Bayesian Analysis

When a trigger occurs, onboard software will determine the probability of the trigger being caused by each event type using a Bayesian analysis. Bayesian analysis is a method of probability in which observed data are compared to how likely it is that a theoretical model would produce the data. The basic Bayesian equation of the probability that a theoretical model A is correct, given observed data when models A through X exist, is

$$\text{Prob}(\text{A} \mid \text{Data}) = \frac{\text{Prior}(\text{A}) * \text{Likelihood}(\text{Data} \mid \text{A})}{\sum_{\text{k}=\text{A}}^{\text{X}} \text{Prior}(\text{k}) * \text{Likelihood}(\text{Data} \mid \text{k})}$$

where Prior(A) is the initial probability that the model was correct before data was observed at all and Likelihood(Data|A) is the probability that the model A would give the observed data. In our case, the theoretical models are the previously described events that are known to cause an increase in gamma radiation. The priors of these events were determined directly from the ratio of BATSE triggers for each model and the total number of BATSE triggers. These priors can be found in Table 1. Even though these priors were obtained from a previous instrument, priors should always be based on actual experience and not theoretical models. After GBM has been in operation for some time, new priors will be calculated based on GBM observations, and the software updated accordingly. Additionally, a very useful characteristic of Bayesian analysis is that the more data that is observed, the less effect the original prior distribution has on the result.

The major task of this project was to determine the likelihood distributions for each type of event. To do this, we compiled BATSE data from triggers corresponding to the different events.

2.3 Parameters

It is important to pick parameters that will effectively distinguish among the event classes. The parameters used in the Bayesian analysis were hardness ratio, McIlwain magnetic L coordinate, geocenter angle, and the right ascension (R.A.) and declination (Dec.) of the event. The hardness ratio was calculated by dividing the strongest trigger count in Channel 2+3 into the trigger count of the same detector in Channel 1. This was a very effective parameter. The vast majority of solar flares and SGRs had a hardness ratio above one, while most of the other events had a ratio below one (see Figure 2). Around one, however, it was not very useful because most of the distribution curves crossed. The McIlwain L coordinate is the number of earth radii out from the center of Earth's magnetic field to the equator of the magnetic field line. A previously determined distribution of particle events over McIlwain L values (Mitrofanov et al. 2004) was used, and helped distinguish particle events from other events. This parameter distinguishes particle events well because many more of them occur at high L values than at low. However, that other events occur more often at low L values is misleading. Other events should have an equal chance of being seen at all L values. This is actually the case, and other events are more likely to be seen at low L values, only because the spacecraft spends much more time at these low values (see Figure 3).

The hardness ratio and maximum over opposite detector count ratio were taken for 115 GRBs, 100 solar flares, 39 magnetospheric particle events, 37 SGRs, 31 instances of X-ray transient J0422+32, and 16 instances of Cygnus X-1. The data was histogrammed, normalized to one, and a smoothing function was applied (see Appendix). Individual datum for McIlwain L coordinate and geocenter angle were not taken because each pertained to only one specific kind of event and were known not to effect the distributions of the other events.

Our program compares the R.A. and Dec. of each event with the locations of the sun and other known sources. The proximity of an event to the sun or another known gamma-emitting source raises the likelihood that it comes from that source. The likelihood distribution used for known sources assumed a standard 2-D Gaussian distribution in location. The error box of the distribution was given by the error in the position of the event. The program uses all of these parameters in the above Bayesian analysis to determine the probability that a trigger is of each event type.

Some precautions had to be taken when gathering the data. Only time periods when BATSE was triggering on Channel 2+3 could be used because hardness ratios were readily available only during these times. When collecting solar flare data, it was found that often a single solar flare caused several consecutive triggers, and only the first trigger was recorded. When gathering data on particle events, all South Atlantic Anomaly events were thrown out. Additionally, rate increases due to rises of known sources such as Cygnus X-1 were not used.

2.4 Tests

We conducted tests to measure the accuracy of the classification algorithm. Independent data, data not used to generate the likelihood distributions, were obtained for the testing of GRBs and solar flares. There wasn't enough BASTE data to obtain independent data for particle events, SGRs, X-ray transients, and Cygnus X-1, so the same data was used for testing as was used to generate the likelihood distributions. For all of the GRB tests, the geocenter angle was set to 90 degrees, the error to 5 degrees, the maximum over opposite ratio to 10, and the sun location to (0, 0). As mentioned above, the maximum over opposite ratio will be used to exclude particle events before the event triggers the program. Therefore, it will not be part of the Bayesian analysis, so we set it to exclude particle events. However, the program will not be able to distinguish distant particle events from GRBs if they are near the geocenter angle of Earth's horizon. There was not enough BATSE distant particle data to find a way to distinguish distant particle events from GRBs other than by their geocenter angle. Distant particle events are assumed to have the same hardness ratio as other particle events, though the hardness ratio distribution of particle events is very similar to the GRB distribution. This is why the geocenter angle was set to 90 degrees for GRB tests, excluding the possibility that the test cases were distant particle events. Since GRBs were shown to occur at random locations across the sky, the exact location in the sky of the sun did not matter. So the location we used, R.A. 0, Dec. 0, was as good as any other. The error of 5 degrees was a reasonable error for BATSE data.

Solar flares were tested with the same parameters as GRBs, except that the actual sun locations were entered in.

In addition to generic X-ray transients, we also tested X-ray transient J0422+32. J0422+32 was tested with the same settings for geocenter angle, error, maximum over opposite ratio, and sun location as the GRBs. We know the location of J0422+32, so the test case's proximity to it would affect its probabilities. Generic X-ray transients were tested the same way, though because there wasn't adequate BATSE data for X-ray transients besides J0422+32, the hardness ratios for the J0422+32 gamma emissions were used. To test generic X-ray transients, we set the location of the event to R.A 100, Dec. 100. This placed the event far from all known sources, so it had the effect of observing a new, unknown transient source.

We also tested a known SGR, SGR1806-20, whose location is known. The same settings for geocenter angle, error, maximum over opposite ratio, and sun location were used as the X-ray transient tests. Generic SGRs were tested with exactly the same settings as SGR1806-20, except the location of SGR1806-20 was removed from the program, so it had the same effect as the generic X-ray transient test.

For the testing of distant particles, the geocenter angle was set to 77 degrees and the error to 5 degrees, so the earth's horizon was within the error bar. For locations, we used the same locations that were used for the GRBs, because they were random. Maximum over opposite ratio was set to ten and sun location to R.A. 0, Dec. 0.

For all test sets, we generated a corresponding set of McIlwain L values that were determined by the distribution calculated by Mitrofanov et al. used in the program. We randomly set these values to the test cases.

2.5 Results

A dynamic table was created to quickly analyze all of the test results in order to determine the best confidence level to use in classifying events. The confidence level is the minimum probability of being an event that a trigger must have to be declared as that event. By simply inputting the desired confidence level, the table shows the results of each event's test with the new confidence level. For each event type, the table shows the percent of the data that was classified as every possible event. An event was said to be unclassified if it could not be classified as an event within the confidence level. Using the estimation that GBM will see 200 GRBs per year and applying the BATSE priors to estimate how many other events GBM will see per year, the table shows the number of GRBs that will be misclassified as another event and the number of other events that will be misclassified as GRBs per year. A good confidence level is the one that produces a low number of events misclassified as GRBs, but also keeps the number of misclassified GRBs low. The best confidence level was found to be around 70%. See table 2 for the test results with the best confidence level.

3. CONCLUSION

The program was very successful with the test data. At the 70% confidence level, it had a 95% success rate classifying GRBs, and an 86% success rate of not identifying other events as GRBs. The majority of events were correctly classified, and those that were misclassified were understood. All GRBs that were misclassified were located near known sources. The few solar flares that were classified as GRBs had hardness ratios on the outskirts of the distribution for flares and their position placed them more than five degrees from the sun. However, each of these misclassified solar flares matched up with solar flares seen by the Geostationary Operational Environmental Satellite (GOES), which is the primary reason the BATSE team classified the events as solar flares. The event that was most misclassified as a GRB was the generic hard X-ray transients. In fact, every such event was classified as a GRB. This is because it is not until these sources produce several triggers that they are distinguishable from GRBs. The same problem is also present, to a lesser extent, with SGRs. The GBM team will still go over each trigger on a daily basis, so these new sources will only be misclassified the first few times they are seen. Once it has been determined that a new source is present, the source will be added to the program's known source list. Subsequent triggers from the new source will now be associated with a known location and will no longer be classified as GRBs. Also, only strong events, which are easier to classify, will initiate a turn of the spacecraft, so several of the events misclassified as GRBs will be ignored for being too weak. However, even though this means less spacecraft time lost, correct classifications are still important for ground based optical and radio searches.

Additionally, having the LAT observe these new sources under the presumption that they are GRBs is not necessarily a bad thing. Getting as much data as possible about a source the first time it is seen could be very useful. Such data could be used to help develop answers to the processes that cause the sources to become active. Because of this, it may be decided that for the initial period after a source is discovered, the LAT should turn to observe when the source triggers the GBM, rather than ignore it.

In conclusion, Bayesian analysis of triggers was effective in classifying the triggers as event types. While it incorrectly classified some events, these events were understood and did not pose a significant problem. It was found that with the best confidence level, only 5.5% of GRBs would be missed, while 13.5% of other events would be misclassified as GRBs. However, in reality, the number of events misclassified as GRBs will be much smaller. The GBM will be less sensitive than BATSE, and thus should trigger less on SGRs and Cygnus X-1, which are weak in the 50-300 keV range. The onboard event classification program was a success and will need only minimal changes to port over to the actual flight version.

References

Gehrels, N. (2002, March 21). *GLAST: Science and Technical Overview*. Retrieved 6 August 2002 from <http://glast.gsfc.nasa.gov/science/ >

Mitrofanov, I.G., Anfimov, D.S., Briggs, M.S., Fishman, G.J., Kippen, R.M., Kozyrev, A.S., Litvak, M.L., Meegan, C.A., Paciesas, W.S., Preece, R.D., & Sanin, A.B. 2004, ApJ, **603**, 624

Event Type	Prior
Gamma Ray Burst	0.54
Generic Soft Gamma Repeater	0.02
Generic Hard X-Ray Transient	0.02
Distant Particles	0.1
Solar Flares	0.24
Cygnus X-1	0.06
SGR1806-20	0.01
GROJ0422+32	0.01

Table 1

70% Confidence	Classified As								
Event	Burst	Gen SGR	Gen X Trans	Particles	SolarFlare	CygX-1	SGR 1806-20	J0422+ 32	Unknown
Burst	0.943	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.057
GenSGR	0.333	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.667
GenXTrans	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Particles	0.103	0.000	0.000	0.641	0.000	0.000	0.000	0.000	0.256
SolarFlare	0.040	0.000	0.000	0.000	0.920	0.000	0.000	0.000	0.040
CygX-1	0.188	0.000	0.000	0.000	0.000	0.563	0.000	0.000	0.250
SGR1806	0.067	0.000	0.000	0.000	0.000	0.000	0.600	0.000	0.333
J0422	0.484	0.000	0.000	0.000	0.000	0.000	0.000	0.290	0.226

Table 2

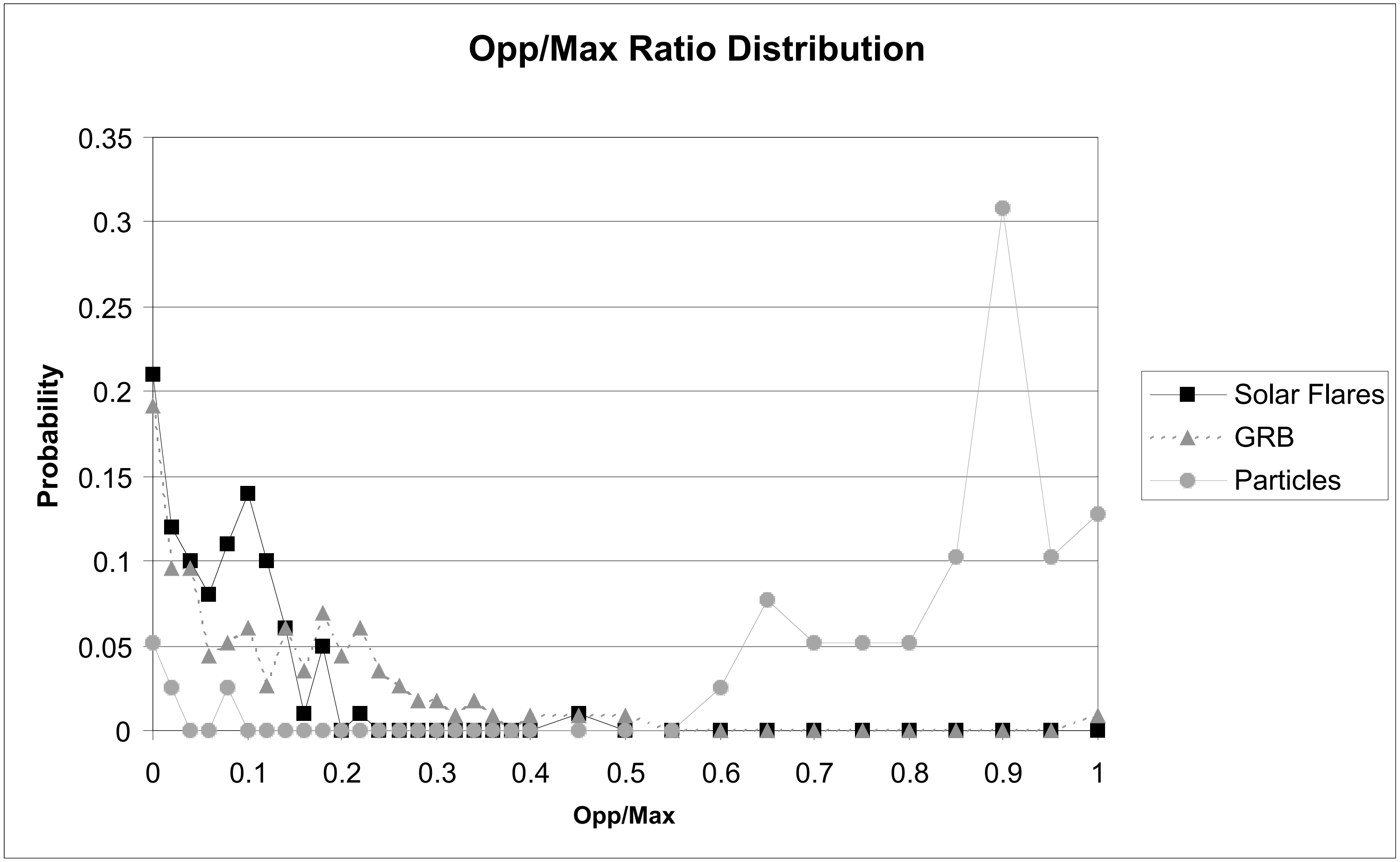
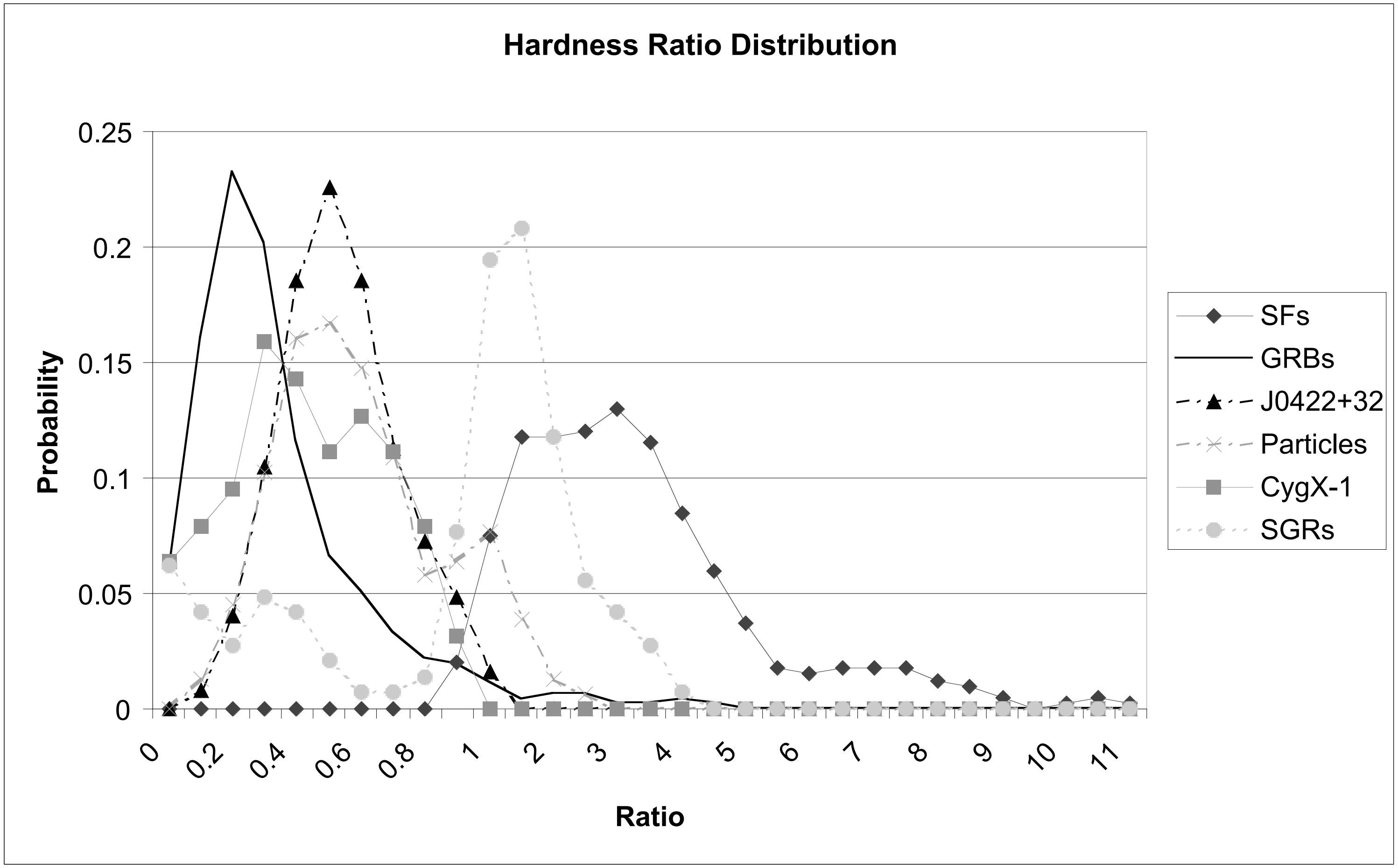


Figure 1

Counts in maximum detector divided by counts in the opposite detector. Because the particles flood all of the detectors equally, the Opp/Max ratio is near one. Other events are only seen in a few detectors, thus the ratio is small.

The hardness ratio was calculated by dividing the counts of the Channel 2+3 into the Channel 1 counts of the same detector. This was an excellent discriminator between GRBs and solar flares. However, it was not very useful with hardness ratios of about one, as all of the distributions cross over each other in that area.


Hardness Ratio Distribution
Probability
Ratio
0.25
0.2
0.15
0.1
0.05
0
0
0.2
0.4
0.6
0.8
1
2
3
4
5
6
7
8
9
10
11
SFs
GRBs
J0422+32
Particles
CygX-1
SGRs

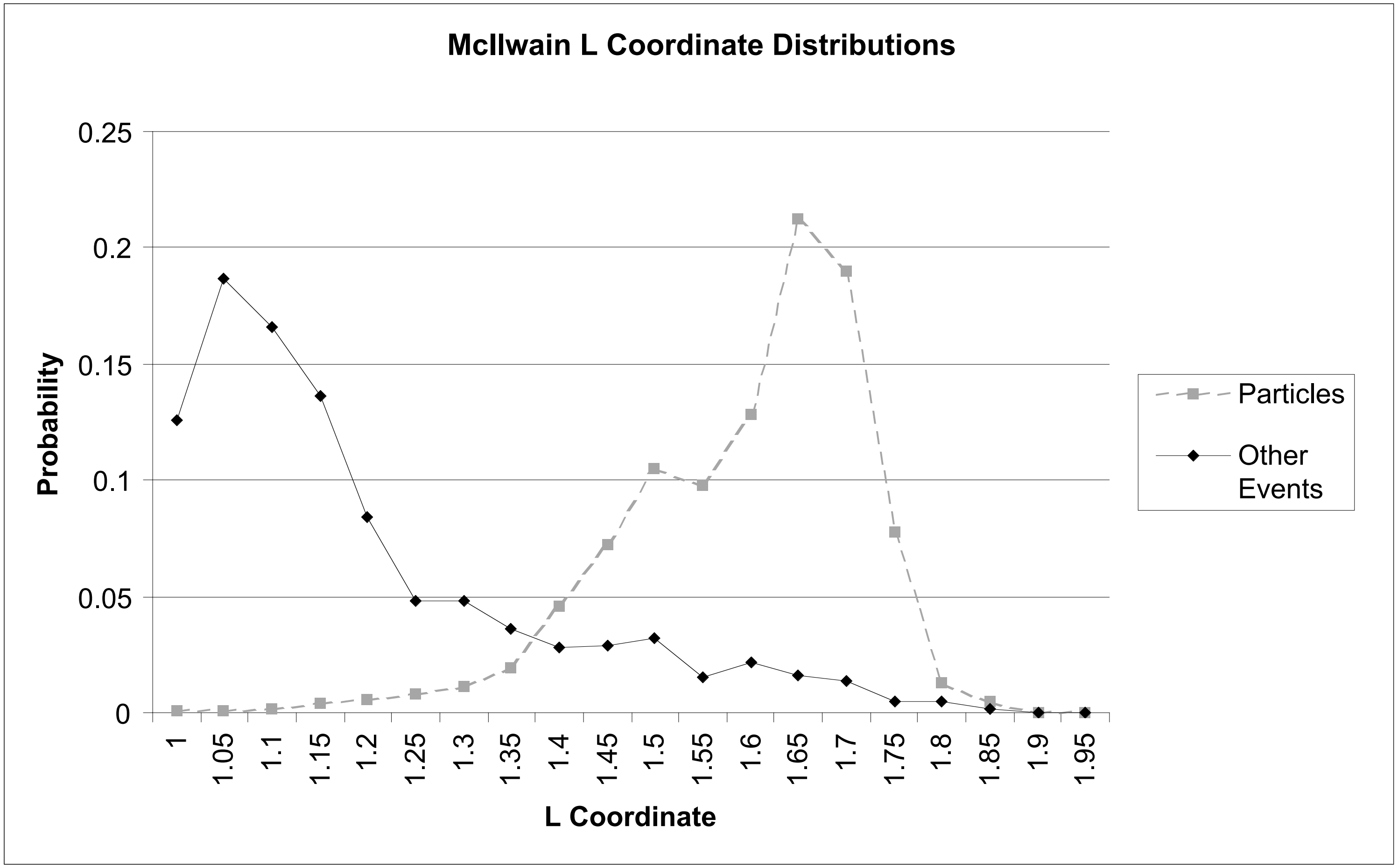


Figure 3

The L coordinate helped distinguish particle events from other events. Particle events were found to occur mostly when the spacecraft was at high L coordinates. The distribution of other events is representative of how much time the spacecraft spent at each L coordinate.

Appendix

A smoothing function was used on the data to give the somewhat jagged data a more Gaussian shape. Additionally, there were some cases in which there were gaps in the data, for example, some GRBs had hardness ratios 1 and 3 and none had a ratio of 2. The smoothing function smoothed out these gaps. The smoothing function for each bin was

New bin value = 0.5*Old bin value + 0.25*value in previous bin
+ 0.25*value in next bin.

The function was applied to each bin using the original bin values. In this way, the function retained the total number of counts in most cases. However, if a data set had values at the edges, a quarter of the edge bin's counts were lost. In these cases, the numbers were renormalized using the new number of counts.