Connecting Swift Long-GRB Populations to the Star Formation Rate

Anjali Mittu

Abstract

Long GRBs could be a probe to conditions of star formation in the early universe since GRBs can be seen in a wide range of redshifts, including high redshifts ($z > 6$). To better measure the SFR using long GRBs, it is important to account for the detection efficiency of Swift as a function of redshift. Swift uses a complex trigger algorithm with more than 500 trigger criteria, thus introducing complicated selection effects on GRBs. Recently in Graff et al (2015), the Swift BAT triggering algorithm was modeled using machine learning algorithms (MLAs) and the detection efficiency of Swift was measured. In this research I reproduced the detection fraction found in Graff et. al. (2015) and found the detection fraction for a flatter luminosity distribution. This was applied in a Bayesian study of the GRB rate distribution. The results show a difference in the observed SFR rate and the long GRB rate.

Introduction

Gamma-ray bursts (GRBs) are some of the most energetic events in the known universe. In just a couple of seconds, GRBs release more than 10^{51} ergs making them the most powerful explosion since the big bang. Because of this, GRBs remain luminous even at high redshift, making them one of the few objects seen at redshifts $>\approx 6$. There are two types of GRBs observed: long GRB and short GRB, which are divided based on their light curves and duration. Short GRBs are thought to come from the merger of two compact objects such as neutron stars, while long GRBs comes from the core collapse of a massive star. Long GRBs have also found to be connected to Type Ic supernovae (Woosley and Bloom 2006).

GRBs are particularly important for studying the star formation rate (SFR). At high redshifts, tracing the SRF is very difficult due to the fact that it is hard to see out so far. This results in estimates of the SFR differing by more than an order of magnitude as seen in figure 1. GRBs should be able to trace the SFR since they can be seen at high redshifts and are the result of the death of a massive star. GRBs could also be useful at low redshifts since they can provide measurements for galaxies normally too dim to see. Previous studies have tried to measure the intrinsic GRB redshift and luminosity distributions and have found that it differs from the SFR (e.g. Yuksel et al. 2008; Kistler et al. 2008; Butler et al. 2010; Wanderman & Piran 2010; Lien et al. 2014; Graff et al. 2015; Graham & Schady 2015).

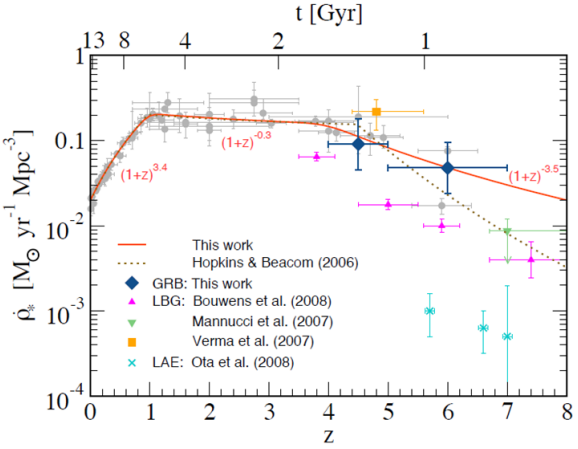

Figure 1: Complied SFR data from Yüksel et al. (2008)

Swift Challenge

The Swift Gamma-Ray Burst Mission (Gehrels et al. 2004) was launched in 2004. It's autonomous rapid-slewing capabilities allow it to quickly detect and localize GRB. This permits ground-based observatories to follow up and measure the GRB's redshift. The burst alert telescope (BAT) on Swift uses a trigger algorithm with $>$ 500 rate triggers. This complex trigger algorithm successfully increases the number of GRB detected, yet makes

estimating the detection fraction more difficult. The complicated trigger introduces complicated selection effects on GRBs. Previous studies estimate the trigger algorithm using a flux detection threshold. This is not a good estimate since the trigger criteria on BAT are based on photon count rates with an addition image trigger (≈every minute) using a real image for missed bursts.

In Lien et. al. (2014) the trigger algorithm was simulated using code that followed the same trigger criteria as BAT. This simulator was able to accurately represent the trigger algorithm. The problem with the trigger simulator was that is took ≈ 10 secs to simulate a single burst. Running the simulator with a large number of burst to cover over a wide distribution of redshift is very computationally intensive. This makes the simulator impractical for very large sets of bursts. Graff et. al. (2015) addressed the problem in the Lien et. al. (2014) simulator by using MLAs to speed up the process. The MLAs were trained and tested using data generated by the Lien et. al. (2014) simulator. The generated data was based of a luminosity distribution function (Eq 1) with constant parameters $x = -0.65$, $y = -3$ and $L_* = 10^{52.05} ergs/s$.

$$\phi(L) = \begin{cases} (\frac{L}{L_*})^x & L \geq L_* \\ (\frac{L}{L_*})^y & L < L_* \end{cases} \quad (1)$$

The results from previous studies for the detection fraction are shown in figure 2.

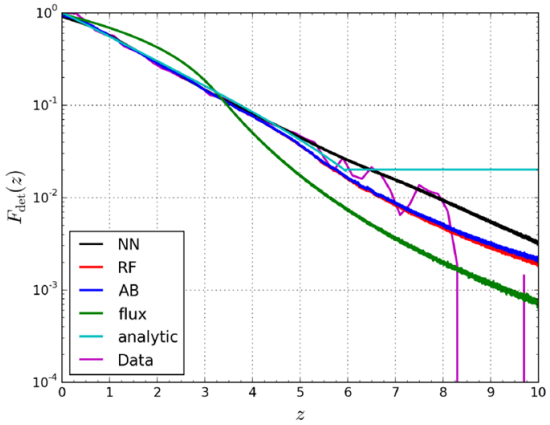

Figure 2: $F_{det}(z_i)$ for three different MLAs from Graff et. al. (2015) with the constant flux cut and analytic form used in Howell et al (2014)

In this study, I reproduced the detection fraction found in Graff et. al. (2015) using random forests (RFs), found the detection fraction for a flatter luminosity distribution and performed a Bayesian analysis on the data using a one-break and two-break model.

Reproducing Detection Fraction

A Random Forest algorithm is an ensemble of multiple decision trees. Each non-leaf node of the decision tree makes a decision based on one of the input variables. Each leaf node is a target value. An import feature of decision trees is that they can create partition recursively (Breiman et. al. 1984; Quinlan & Ross 1993). With each split the tree is making a decision based on a smaller subset of data then the one before. RFs improves on this method by averaging multiple decision trees together. To create multiple trees from one dataset you use a technique called the bootstrap. This means that each tree uses a random subset of the data. Another trick used in Random Forests is to use a random selection of features to split when deciding partitions. This reduces similarities between trees.

$log_{10}(L)$	luminosity of the GRB
z	redshift
r	distance from cnter of detection grid of peak
ϕ	azimuthal angle in the detector grid of peak
bin_size_emit	source time bin size
α	Band function parameter
β	Band function parameter
$log_{10}(E_{peak})$	peak of energy spectrum of the GRB
bgd_15_25eV	background count rate in 15-25eV band
bgd_15_50eV	background count rate in 15-50eV band
bgd_25_100eV	background count rate in 25-100eV band
bgd_50_350eV	background count rate in 50-350eV band
θ	incoming angle of GRB
$log_{10}(\Phi)$	incident flux of GRB
$ndet$	number of active detector pixels (constant)
$trigger\_index$	0 for non-detection and 1 for detection

Table 1

A set of 15 parameters were used as inputs to the algorithm (Table 1). I tried multiple data sets with various number of samples to produce a similar detection fraction as in Graff et. al. (2015). These data sets came from the data generated for Lien et. al. (2014). For each data set, $\frac{4}{5}$ of the data was used for training with the remaining $\frac{1}{5}$ left out for testing. Additionally each model was then tested against 5 separate data sets each with ≈ 10000 samples. A summary of these data sets and their test scores are shown in table 2.

Name of Data Set	Number of Samples	Accuracy against training set	Accuracy against test set
Set 1a	4100	98.4%	94.4%
Set 1b	5900	99.2%	94.0%
Set 1 Combined	10000	99.1%	95.1%
Set 2	10000	99.2%	94.9%
Set 1 and 2 Combined	20000	99.4%	96.2%

Table 2

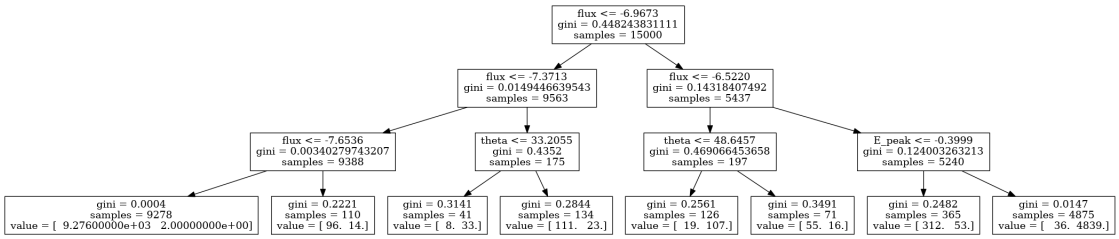

Figure 3: An example decision tree from the Set 1 and 2 Combined RF model. Each non-leaf node shows the parameter used for branching, the Gini factor, and the number of samples used to make the decision. Each leaf node shows the Gini factor, and the number of samples ending at that node, and the number of samples that were detected or not.

The detection fraction $F_{det}(z)$ is computed by using the ML model to predict the percentage of bursts measured at certain redshifts. Between the redshifts of $[0, 10]$, 10^5 GRB are simulated at 1001 evenly spaced redshifts. These points are then used for a spline interpolation to get $F_{det}(z)$. A comparison of the detection fraction produced by the data sets is shown in Figure 4

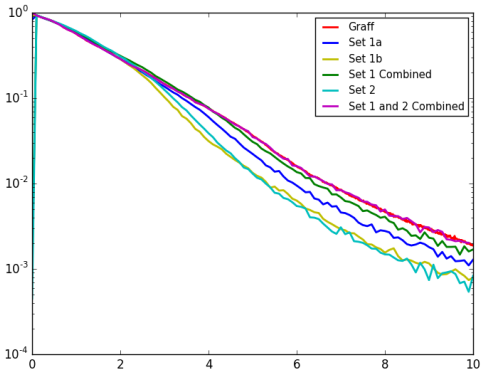

Figure 4: The detection fraction from Graff et. al. (2015) along with the five from the new RF models.

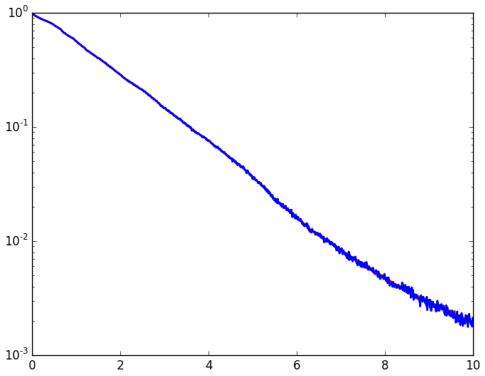

Figure 5: The detection fraction from Graff et. al. (2015) (red) with the best matching RF model (blue).

From these results you can see that the Set 1 and 2 combined best matched the detection fraction as in Graff et. al. (2015) and was also the most accurate at predicting. Since this model was the closest to the Graff et. al. (2015) detection fraction, the detection fraction was recalculated using 10^5 GRB are simulated at 100001 evenly spaced redshifts to get a more accurate function. This is shown in Figure 5. An example decision tree with a maximum depth of three from this MF model is shown in Figure 3.

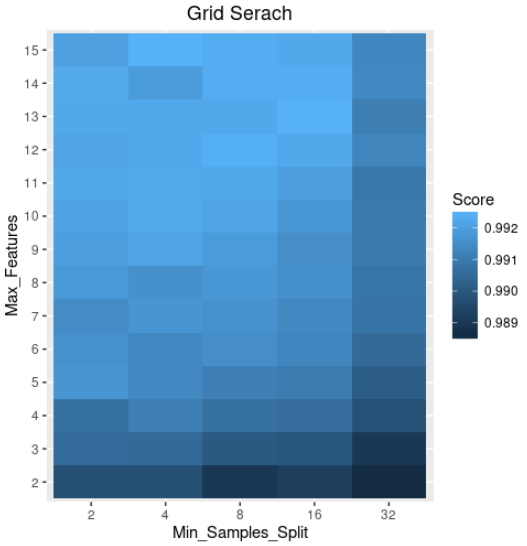

Figure 6: Grid search results for Set 1 and 2 Combined RF.

The hyper-parameters for all the the RF models produced above were optimized using a 5-fold cross validation grid search. The choices for hyper-parameters are as follows:

$$min\_samples\_split \in [2, 4, 8, 16, 32]$$

$$max\_features \in [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15]$$

The min_samples_split is the minimum number of samples required at a node to split the node. The max_features parameter is the maximum number of features to consider when looking for the best

split. The results of the grid search are shown in Figure 6. The values of the best fit model are $min\_samples\_split = 2$ and $max\_features = 11$.

New Detection Fraction

The next step was to build a RF model based on a flatter luminosity distribution

$$\phi(L) = \begin{cases} (\frac{L}{L_*})^{\sqrt{x}} & L \geq L_* \\ (\frac{L}{L_*})^{\sqrt{y}} & L < L_* \end{cases} \quad (2)$$

The same 15 parameters from Table 1 were used as inputs to train the next set of RF models. I used a date set with 10,000 samples where $\frac{1}{5}$ was left out for testing. The accuracy of this model against the training set is 95%. This is much lower than the previous models.

The hyper-parameters for this models were also optimized using a 5-fold cross validation grid search. The choices for hyper-parameters are as follows were the same as above. The results of the grid search are shown in Figure 7. The values of the best fit model are $min\_samples\_split = 16$ and $max\_features = 13$. The difference between the best values and the worst values was only .5% which was similar to the previous grid search. However this search did not show a correlation between $min\_samples\_split$ and $max\_features$ as the previous grid search did.

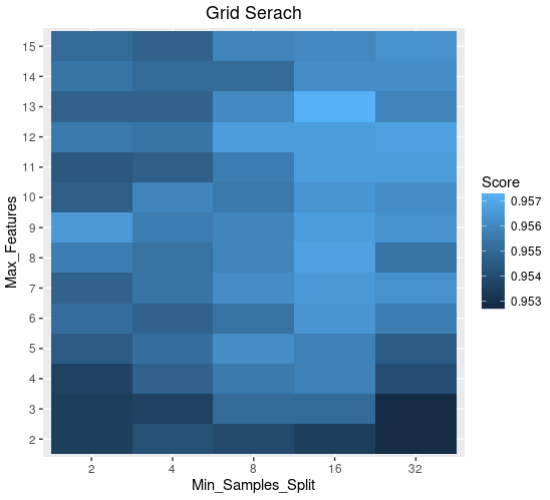

Figure 7: Grid search results for variable luminosity distribution RF.

The detection fraction was computed by simulating 10^5 GRB at 1001 evenly spaced redshifts. A comparison of this detection fraction and the previous detection fraction is shown in Figure 8. The detection fraction found with the flatter luminosity was higher then before.

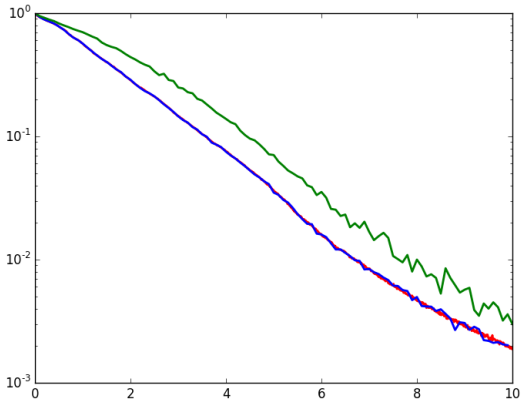

Figure 8: The detection fraction from Graff et. al. (2015) with the best matching RF model. The function from Graff et. al. (2015) is shown in red, my reproduced function shown in blue, and the flatter luminosity function shown in green.

Analysis of GRB

We can also fit a model to gain information about the rate of GRB. The number of intrinsic GRBs occurring in Swift's field of view is given by Equation 1.

$$N_{int}(z_i) = \frac{4\pi}{6} \Delta t_{obs} R_{GRB;dz}(z_i) dz \quad (3)$$

Where the factor of $4\pi/6$ accounts that Swift can only observe a sixth of the sky at any time. $\Delta t = .8$ which is the amount of time Swift spends observing per year. $R_{GRB;dz}(z)$ is the comoving rate accounting for time dilation and comoving volume.

The expected number of observed GRBs comes from multiplying Equation 1 by the detection fraction.

$$N_{exp}(z_i) = \frac{4\pi}{6} \Delta t_{obs} R_{GRB;dz}(z_i) F_{det}(z_i) dz \quad (4)$$

Using Equation 1 and 2, the log-likelihood is found to be Equation 3, where $\{i\}_{det}$ is each detection.

$$\mathcal{L}(n) = -N_{exp} + \sum_{\{i\}_{det}} log(N_{exp}(z_i)) \quad (5)$$

The full deviation of the log-likelihood can be found in Graff et. al. (2015).

One-break vs. Two-break

In this analysis I test two models for $R_{GRB}(z)$ to find find the best fit: a one-break model shown in Equation 6 and a two-break model shown in Equation 7. The variables n_i and z_i used in Equation 6 and Equation 7 were allowed to vary.

$$R_{GRB} = \begin{cases} (1+z)^{n_1} & z \leq z_1 \\ (1+z)^{n_1-n_2}(1+z)^{n_2} & z > z_1 \end{cases} \quad (6)$$

$$R_{GRB} = \begin{cases} (1+z)^{n_1} & z \leq z_1 \\ (1+z)^{n_1-n_2}(1+z)^{n_2} & z_1 < z \leq z_2 \\ (1+z)^{n_2-n_3}(1+z)^{n_1-n_2}(1+z)^{n_2} & z > z_2 \end{cases} \quad (7)$$

The BAMBI algorithm (Graff et al. 2012) was used to perform the Bayesian parameter estimates on the variables. This uses a real data samples of 66 GRBs from the Fynbo et.al. (2009) study to fit the model and the detection fraction found from my random forest model.

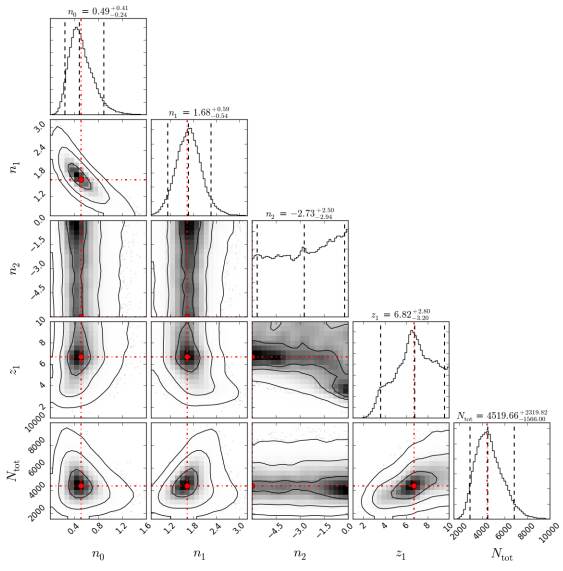

Figure 9: Posterior distribution for the 66 GRBs from Swift using random forest for the one-break model. The red dashed lines show the maximum likelihood value. Contour lines are shown every 68%, 95%, and 99%. Dashed lines are show 5%, 50% and 95% quantiles. N_{tot} is the predicted total number of GRB in the universe per year.

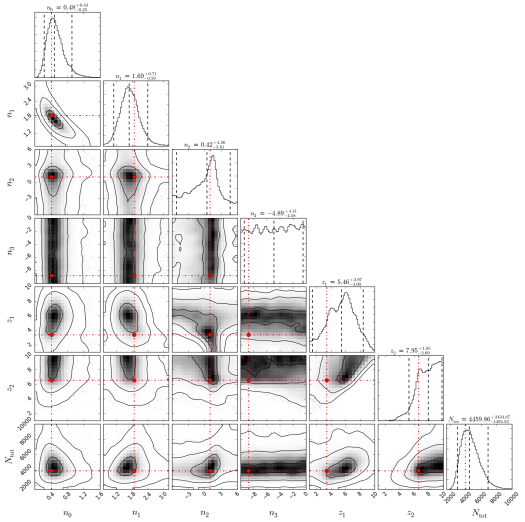

Figure 10: Posterior distribution for the 66 GRBs from Swift using random forest for the two-break model. This is similar to Figure 9

The posterior distribution of the parameters are shown in Figure 9 and Figure 10. The best fit parameters for the one-break model is $\{n_0, n_1, n_2, z_1, N_{tot}\} = \{.49, 1.68, -2.73, 6.82, 4520\}$ and for the two-break model $\{n_0, n_1, n_2, n_3, z_1, z_2, N_{tot}\} = \{.48, 1.69, .42, -4.89, 5.46, 7.95, 4460\}$. Both models found similar values for n_0, n_1 and N_{tot} but had very different values for n_2. This is because neither model was able to constrain the high redshift parameters to a highly probably value.

The distribution of models for are are shown in Figure 11 and Figure 12. These plots show a random selection of 200 models selected from the posterior as light blue lines, with the most likely values in black. These plots also show that the model is not well constrained at high redshifts.

The Bayesian evidence value for the one-break model was 99.43 +/- 0.0441 and the two break was 99.23 +/- 0.0450. The Bayesian analysis does not find preference for either model, and thus does not support the additional complexity of the two-break model, which tends to mimic the results of the one break model.

Model Proportional to SFR

I ran a two-break model proportional to the SFR as found in Hopkins and Beacom (2008) by using their parameters $\{n_1, n_2, n_3, z_1, z_2\} = \{3.28, 0.26, 8.0, 1.04, 4.48\}$ and varying n_0. The results of this analysis are shown in Figure 13. The model was again compared to the Fynbo et.al. (2009) data, and used the RF detection fraction. The best fit parameter for n_0 is .33.

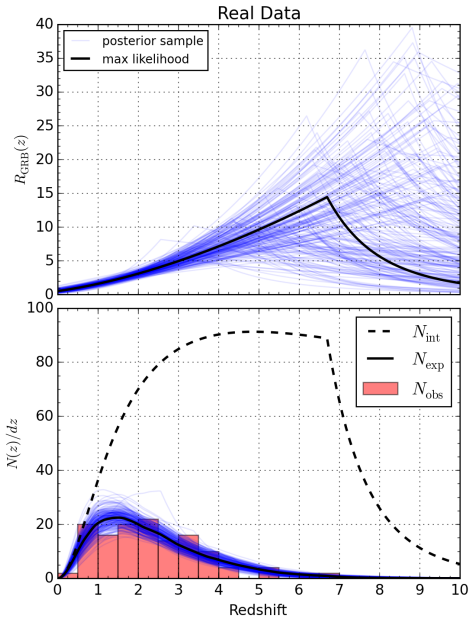

Figure 11: Distribution of model parameters for one-break model. The top plot shows $R_{GRB}(z)$. The light blue line shows a random selection of 200 models selected from the posterior. The black line shows the maximum likelihood points. The bottom plot shows N_{exp}/dz. The red boxes are the GRB selected from the Fynbo study.

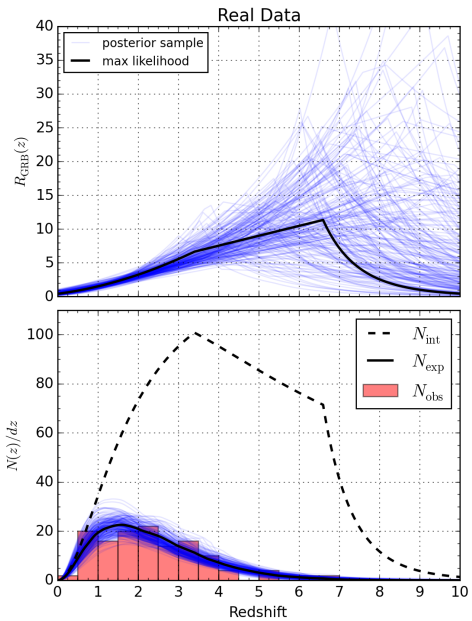

Figure 12: Distribution of model parameters for two-break model. This is similar to Figure 11

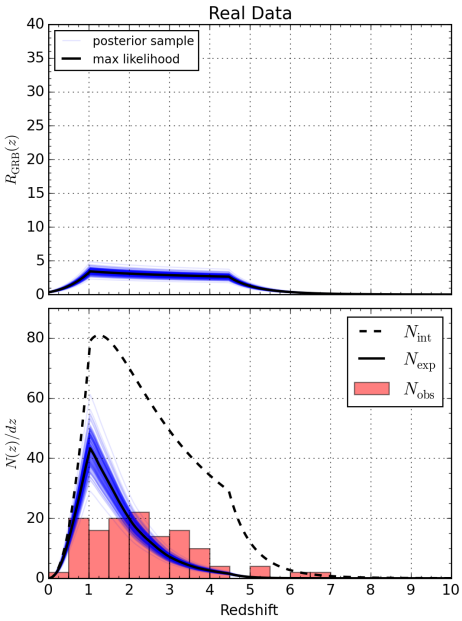

Figure 13: Distribution of model parameters for SFR proportional two-break model. This is similar to Figure 11

The Bayesian evidence of this model was far less than my one- break or two-break; 84.22 vs. 99.43 and 99.23. In this case the Bayesian evidence does show a strong preference for the one-break model over the model proportional to the SFR. Figure 13 shows that the predicted number of detected GRB does not fit the amount of observed. With a model proportional to SFR, one would expect to see many more GRB at redshifts $\approx 1-2$ and less GRB at redshifts around 3.5.

Model with Updated GRB

In both the one-break and two-break model the high redshift parameters were hard to constrain. The Fynbo et al (2009) dataset comprised of 66 GRB, with only two at high redshifts. This provided little constraint on the higher redshift population $z > 4$. In the past seven years since the Fynbo sample there have been hundreds more GRBs detected. The SHOALS (Perley et al 2016) sample contains 112 GRB and can be used to incorporate more recent data. The results of fitting the SHOALS sample are shown in Figure 14. The best fit parameters for the one-break model is $\{n_0, n_1, n_2, z_1, N_{tot}\} = \{1.18, 1.37, -2.95, 6.01, 6190\}$. Even with the additional high redshift GRB, the high redshift parameter n_2 was still not able to be constrained. These values are graphed in Figure 15. In this Figure you can see that the gap between predicted observed GRB and real observed GRB increased

at a redshift of 1.5 from the Fynbo model.

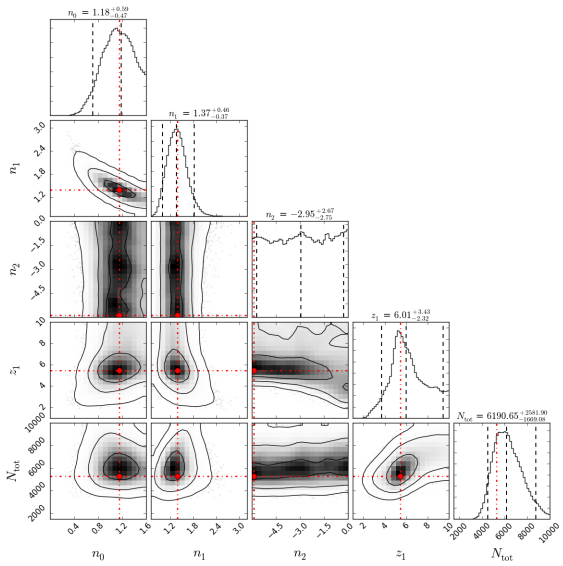

Figure 14: Posterior distribution one-break model based on SHOALS sample (Perley et al 2016). This is similar to Figure 9

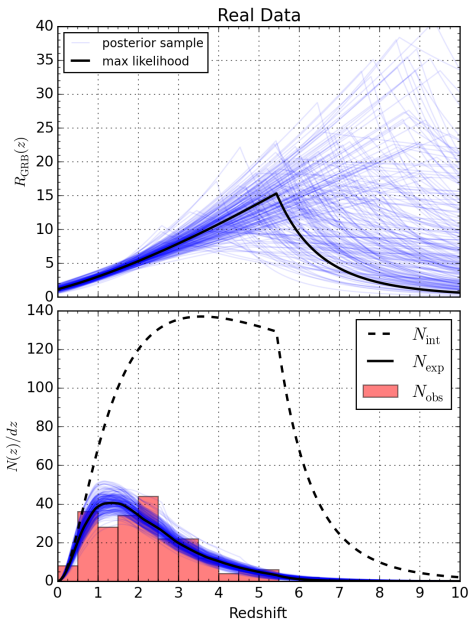

Figure 15: Distribution of model parameters for one-break model. This is similar to Figure 11 except that the red boxes come from the SHOALS sample (Perley et al 2016).

To compare this model with the Fynbo model, n_0 must be normalized for the same about of time. Swift observes about 100 GRBs per year. This means that the Fynbo sample contains less than a year's worth of GRBs (66 GRBs) and the SHOALS sample has more than a year's worth (112 GRBs). For the Fynbo sample, n_0 must be multiplied by $\frac{100}{66}$ which gives .74. For the SHOALS sample, n_0 must be multiplied by $\frac{100}{112}$ which gives 1.05. n_1, n_2, and z_1 are not affected since they don't need to be normalized for a length of time. A comparison of all the models is shown in Figure 16.

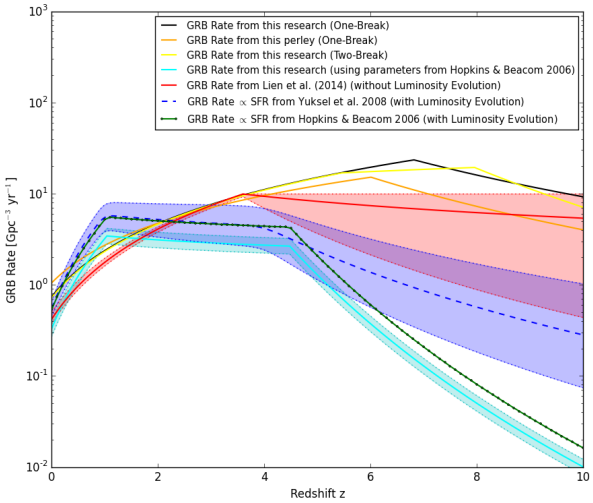

Figure 16: Comparison of GRB rate models.

While the SHOALS model wasn't able to give more insight to higher redshift rates, it did provide an interesting observation for lower redshifts. The SHOALS model predicted a higher number of GRB at redshifts around 1-2 than the Fynbo model. This number came closer to the predictions from the SFR. Both models agreed on the rate of GRB around a redshift of 3. This agreed rate was higher than the SFR prediction.

References

L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone. Classification and Regression Trees, Wadsworth, Belmont, CA, 1984. Since 1993 this book has been published by Chapman & Hall, New York.

N. R. Butler, J. S. Bloom, and D. Poznanski. The Cosmic Rate, Luminosity Function, and Intrinsic Correlations of Long Gamma-Ray Bursts. ApJ, 711:495–516, March 2010. doi:10.1088/0004-637X/711/1/495.

J. P. U. Fynbo, et. al. Low-resolution Spectroscopy of Gamma-ray Burst Optical Afterglows: Biases in the Swift Sample and Characterization of the Absorbers. ApJS, 185:526–573, December 2009. doi:10.1088/0067-0049/185/2/526

N. Gehrels et. al. The Swift Gamma-Ray Burst Mission. ApJ, 611:1005–1020, August 2004. doi:10.1086/422091

J. F. Graham, P.Schady. The Absolute Rate of LGRB Formation. doi:10.3847/0004-637X/823/2/154

P. Graff, F. Feroz, M. P. Hobson, A. Lasenby. BAMBI: blind accelerated multimodal Bayesian inference. MNRAS, 421:169–180, March 2012. doi:10.1111/j.1365-2966.2011.20288.x.

P.B. Graff, A.Y. Lien, J.G. Baker, T. Sakamoto, 2015. Machine Learning Model of the Swift/BAT Trigger Algorithm for Long GRB Population Studies. arXiv preprint arXiv:1509.01228.

M. D. Kistler, H. Yuksel, J. F. Beacom, and K. Z. Stanek. An Unexpectedly Swift Rise in the Gamma-Ray Burst Rate. ApJ, 673:L119–L122, February 2008. doi:10.1086/527671.

A. Lien, T. Sakamoto, N. Gehrels, D. M. Palmer, S. D. Barthelmy, C. Graziani, and J. K. Cannizzo. Probing the Cosmic Gamma-Ray Burst Rate with Trigger Simulations of the Swift Burst Alert Telescope. ApJ, 783:24, March 2014. doi:10.1088/0004-637X/783/1/24.

Quinlan, J. Ross. "C4. 5: Programming for machine learning." Morgan Kauffmann (1993): 38.

D. Wanderman and T. Piran. The luminosity function and the rate of Swift's gamma-ray bursts. MNRAS, 406:1944–1958, August 2010. doi:10.1111/j.1365-2966.2010.16787.x

S.E. Woosley, J.S. Bloom. The Supernova–Gamma-Ray Burst Connection. doi: 10.1146/annurev.astro.43.072103.150558

H. Yuksel, M. D. Kistler, J. F. Beacom, and A. M. Hopkins. Revealing the High-Redshift Star Formation Rate with Gamma-Ray Bursts. ApJ, 683: L5–L8, August 2008. doi:10.1086/591449.