The Astrophysical Journal, 756:44 (13pp), 2012 September 1 doi:10.1088/0004-637X/756/1/44

ON THE SPECTRAL LAGS AND PEAK COUNTS OF THE GAMMA-RAY BURSTS DETECTED BY THE *RHESSI* SATELLITE

J. Řípa[1,2,7], A. Mészáros[1], P. Veres[3,4,5,6], and I. H. Park[2,7]
[1] Faculty of Mathematics and Physics, Astronomical Institute, Charles University, V Holešovičkách 2, CZ 180 00 Prague 8, Czech Republic; ripa@sirrah.troja.mff.cuni.cz, meszaros@cesnet.cz
[2] Institute for the Early Universe, Ewha Womans University, 11-1 Daehyun-dong, Seoul 120-750, Republic of Korea
[3] Department of Astronomy and Astrophysics, Pennsylvania State University, 525 Davey Lab, University Park, PA 16802, USA; veresp@psu.edu
[4] Konkoly Observatory, POB 67, H-1505 Budapest, Hungary
[5] Department of Physics of Complex Systems, Eötvös University, Pázmány P. s. 1/A, H-1117 Budapest, Hungary
[6] Department of Physics, Bolyai Military University, POB 15, H-1581 Budapest, Hungary
[7] Department of Physics and Research Center of MEMS Space Telescope, Ewha Womans University, 11-1 Daehyun-dong, Seoul 120-750, Republic of Korea; ipark@ewha.ac.kr
Received 2011 August 8; accepted 2012 June 25; published 2012 August 13

ABSTRACT

A sample of 427 gamma-ray bursts (GRBs) from a database (2002 February–2008 April) of the *RHESSI* satellite is analyzed statistically. The spectral lags and peak-count rates, which have been calculated for the first time in this paper, are studied, completing an earlier analysis of durations and hardness ratios. The analysis of the *RHESSI* database has already inferred the existence of a third group with intermediate duration, apart from the so-called short and long groups. The first aim of this article is to discuss the properties of these intermediate-duration bursts in terms of peak-count rates and spectral lags. The second aim is to discuss the number of GRB groups using another statistical method and by also employing the peak-count rates and spectral lags. The standard parametric (model-based clustering) and non-parametric (K-means clustering) statistical tests together with the Kolmogorov–Smirnov and Anderson–Darling tests are used. Two new results are obtained. (1) The intermediate-duration group has properties similar to those of the group of short bursts. Intermediate and long groups appear to be different. (2) The intermediate-duration GRBs in the *RHESSI* and *Swift* databases seem to be represented by different phenomena.

Key word: gamma-ray burst: general

Online-only material: machine-readable table

1. INTRODUCTION

Mazets et al. (1981), Norris et al. (1984), Kouveliotou et al. (1993), and Aptekar et al. (1998) have suggested the division of gamma-ray bursts (GRBs) into two categories, short and long, according to their duration (at ∼2 s). Many observations demonstrate different properties of short and long bursts. They have different redshift distributions (Bagoly et al. 2006; O'Shaughnessy et al. 2008) and may have different celestial distributions (Balázs et al. 1998, 1999; Mészáros et al. 2000; Litvin et al. 2001; Mészáros & Štoček 2003; Vavrek et al. 2008). At present, the predominant opinion is that they are physically different phenomena (Norris et al. 2001; Balázs et al. 2003; Fox et al. 2005; Kann et al. 2011).

There are also statistical indications of a third "intermediate" group. The division of GRBs into three groups has been studied statistically over different databases: BATSE (Horváth 1998, 2002; Mukherjee et al. 1998; Balastegui et al. 2001; Horváth et al. 2006; Chattopadhyay et al. 2007), *BeppoSAX* (Horváth 2009), *Swift* (Horváth et al. 2008, 2010; Huja et al. 2009; Veres et al. 2010), and *Ramaty High Energy Solar Spectroscopic Imager* (*RHESSI*; Řípa et al. 2009). These three groups may also have different celestial distributions (Mészáros et al. 2000; Vavrek et al. 2008), at least for the BATSE database. No test has given statistically significant support for the existence of four or more groups. Only the BATSE database gave a weak 6.2% significance level for such a possibility (Horváth et al. 2006).

One cannot exclude an eventuality that the separation of this third group is simply a selection effect (Hakkila et al. 2000; Rajaniemi & Mähönen 2002). In other words, a separation from the statistical point of view does not necessarily indicate astrophysically different phenomena. In principle, it is still possible that the class of intermediate GRBs constitutes a "tail" of either the short or the long group. The article by Veres et al. (2010) claims that—at least for the *Swift* database (Sakamoto et al. 2008)—the third group is related to the so-called X-ray flashes (XRFs), which need not be physically distinct phenomena (Kippen et al. 2003; Soderberg et al. 2006). Two models of XRFs are favored; they are either ordinary long GRBs viewed slightly off-axis (Zhang & Mészáros 2002) or intrinsically soft long-duration GRBs (Gendre et al. 2007). Hence, at least in the *Swift* database, the problem of the intermediate class seems to have been solved.

However, for three reasons the situation has not yet been clarified. First, with regard to the *Swift* database, another study suggests that even the short group should be further separated (Sakamoto & Gehrels 2009). Second, there is additional observational evidence against the simple scheme that maintains the existence of *only* two types of bursts (short/hard and long/soft) separated at duration of ∼2 s: the GRB 060614 event, which is clearly long at duration (≃100 s) but in any other properties resembles a short GRB; and subsequent short bursts with soft extended emission also challenged this scheme (Gehrels et al. 2006). To avoid the limitations of a short–long separation terminology, the designations "Type I" and "Type II" have been proposed (Zhang 2006; Zhang et al. 2009; Kann et al. 2011) because duration alone is hardly sufficient for a correct division into categories. Third, it remains possible that in other databases the discovered intermediate group is not represented by XRFs. Concerning this third reason, the mean duration of the

intermediate group appears to vary according to the database in which it is found. For the *Swift* data (Horváth et al. 2008; Huja et al. 2009) the mean duration is ∼12 s, which resembles the durations of the long GRBs, but for the *RHESSI* and BATSE data (Horváth 1998; Mukherjee et al. 1998; Horváth et al. 2006; Řípa et al. 2009) this mean is far below 10 s.

It is clear that any new result to aid in the classification scheme of GRB groups is desirable. In this article we study the *RHESSI* database, where in addition to Řípa et al. (2009), the spectral lags and peak counts are also included. We have two concrete aims here: first, to provide further statistical tests concerning the GRB classes and, second, to provide additional information concerning the physical significance of the *RHESSI* intermediate group found by Řípa et al. (2009).

The paper is organized as follows: In Section 2, the *RHESSI* satellite and its GRB data sample are described. In Section 3, distributions of spectral lags, normalized lags, and peak-count rates are studied using Kolmogorov–Smirnov (K-S) and Anderson–Darling (A-D) tests along with Monte Carlo (MC) simulations. In Section 4, we discuss the results of these tests, compare the results with the BATSE and *Swift* data samples, and discuss the number of GRB groups using model-based and K-means clustering methods. Section 5 summarizes the results of this paper.

2. THE *RHESSI* DATA SAMPLE

RHESSI[8,9] is a satellite designed for the observation of hard X-rays and gamma rays from solar flares (Lin et al. 2002), but it is also able to detect GRBs. There is no automatic search routine for GRBs, and the *RHESSI* data are searched for a GRB signal only if a message from any other instruments of the International Planetary Network occurs. Therefore, our data set includes only events confirmed by other satellites.

In this paper, we study the same list of bursts that was published in Řípa et al. (2009). We consider 427 GRBs from the period between 2002 February 14 and 2008 April 25. Contrary to Řípa et al. (2009), the spectral lags and the peak counts—calculated for the first time for *RHESSI*—in addition to the durations and hardnesses are used. They are collected, together with their uncertainties, in Table 9. These new observational data allow further study of the questions concerning the GRB classification. There are two arguments for the choice of the same list of bursts. First, both in Řípa et al. (2009) and in the present work, similar statistical studies are performed. Hence, for comparison, it is reasonable to study the structure of groups found in the *RHESSI* database over the same set. The second argument concerns an instrumental effect. The measurements of the hardness ratio of the events during the year 2008 and later have been systematically affected by an "annealing" procedure[10] executed on the *RHESSI* detectors in late 2007 (Bellm et al. 2008). The *RHESSI* team decided to anneal the detectors to recover its deteriorating spectral sensitivity. However, the sensitivity at low energies was not recovered as well as that at high energies; hence, the measured GRB hardness ratios from the post-annealing period are systematically shifted to higher values (Veres et al. 2009; Řípa et al. 2010). In order to eliminate this instrumental influence, more sophisticated modeling is required. However, this is beyond the scope of this article.

[8] http://hesperia.gsfc.nasa.gov/hessi
[9] http://grb.web.psi.ch
[10] http://hesperia.gsfc.nasa.gov/hessi/news/jan_16_08.htm

Table 1
Six *RHESSI* GRBs with Corrected T_{90} Durations and Hardness Ratios

GRB[a]	Peak Time[b]	T_{90}[c] (s)	Hardness Ratio, log H[d]
030518B	03:12:23.050	(1.86 ± 0.07)E+1	(2.90 ± 0.27)E-1
030519A	09:32:22.500	(3.20 ± 0.27)E+0	(5.31 ± 0.61)E-1
031024	09:24:14.350	(4.30 ± 0.17)E+0	−(2.06 ± 0.31)E-1
040220	00:55:15.800	(1.80 ± 0.07)E+1	(9.39 ± 2.72)E-2
050216	07:26:34.275	(4.50 ± 0.56)E-1	(2.33 ± 0.48)E-1
050530	04:44:44.900	(2.40 ± 0.26)E+0	(2.41 ± 0.63)E-1

Notes.
[a] *RHESSI* GRB number.
[b] Peak time of the count light curve in UTC.
[c] The uncertainties were calculated though the same procedure used in Řípa et al. (2009).
[d] The hardness ratio was defined as the ratio of GRB counts at two different bands, $H = S_{(120-1500)\,\mathrm{keV}}/S_{(25-120)\,\mathrm{keV}}$.

In order to compare the spectral lags and the peak counts of bursts belonging to the different groups, one must provide a rule by which the particular GRBs are sorted into the concrete groups. We proceeded in the following manner. The probability density function employed in the fitting of the duration–hardness plane in Řípa et al. (2009) is composed from the summation of three bivariate lognormal functions, $f(x, y) = f_1(x, y) + f_2(x, y) + f_3(x, y)$, where x is the base 10 logarithm of the duration and y is the base 10 logarithm of the hardness ratio; f_1, f_2, and f_3 are components corresponding to the particular groups. A burst at the point $[x_0; y_0]$ is considered short, intermediate, or long depending on whether $f_1(x_0, y_0)$, $f_2(x_0, y_0)$, or $f_3(x_0, y_0)$ is maximal. In essence, we follow a procedure identical to that of Horváth et al. (2006, 2010) utilized on the BATSE and *Swift* data sets.

In order to sort the given GRBs into the groups, we employ the measurements of the durations and hardness ratios as given in Table 7 of Řípa et al. (2009) with the exception of six events. We find that for these six events the values mentioned in Řípa et al. (2009) were not corrected for a so-called decimation, which is an instrumental mode used to conserve the onboard memory. Table 1 presents these six events, now corrected for this decimation.

The group members used in this study were determined from the best maximum likelihood (ML) fit (Řípa et al. 2009) in the duration–hardness plane of 427 GRBs. In this sample, the six events with corrected decimation were included along with the remaining 421 events taken from Řípa et al. (2009). The best ML fit with two bivariate lognormal components gives logarithmic likelihood $\ln L_2 = -313.4$. The best fit with three components gives logarithmic likelihood $\ln L_3 = -303.4$. The ML ratio test tells us that the difference in the logarithmic likelihoods multiplied by 2, i.e., $2(\ln L_3 - \ln L_2) = 20.0$, should follow a χ^2 distribution with six degrees of freedom (Horváth et al. 2006). Therefore, the ML ratio test, employed in Řípa et al. (2009) and now applied on the duration–hardness plane with these 427 GRBs including the six events corrected for decimation, again gives a statistically significant intermediate group at the significance level of 0.3%. The new (former) best-fit model parameters of the intermediate group are 0.12 (0.11) for the mean logarithmic duration, 0.25 (0.27) for the mean logarithmic hardness, 4.1% (5.3%) for the weight, and 0.0 (0.59) for the correlation coefficient. The group members are shown in Figure 1 and listed in Table 9.

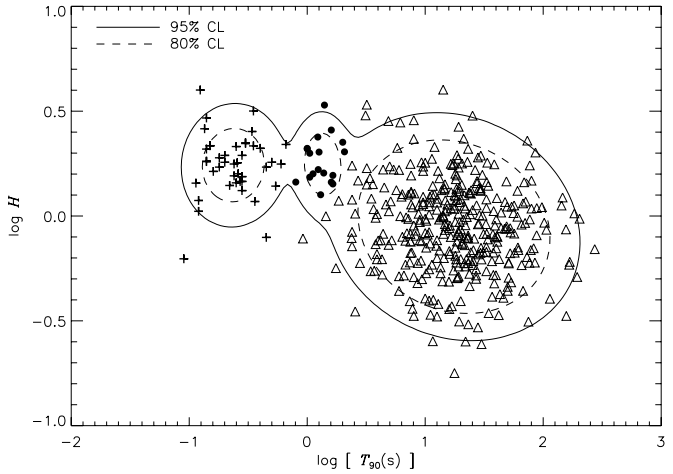

Figure 1. Hardness ratio H plotted against the duration T_{90} for the *RHESSI* database with the best ML fit of three bivariate lognormal functions. The different GRB group members are denoted with different symbols: crosses, full circles, and triangles correspond, respectively, to the short, intermediate, and long bursts. CL means "confidence level."

The spectral lags L of the *RHESSI* data were calculated by fitting the peak of the cross-correlation function (CCF) of the background-subtracted count light curves at two channels, 400–1500 keV and 25–120 keV, by a third-order polynomial. The position of the maximum of the polynomial fit measures the spectral lag. An example of such a fit is shown in Figure 2. The method is similar to that employed in previous studies (Norris et al. 2000; Norris 2002; Foley et al. 2008, 2009) on the BATSE and *INTEGRAL* data. This is the first time that the spectral lags have been calculated for the *RHESSI* GRBs.

To obtain statistical errors, an MC method was utilized. The following procedure was employed to prepare 1001 synthetic count profiles for each GRB. The measured count profiles were randomly influenced by Poisson noise, after which the background was subtracted. The *RHESSI* count rates are sometimes "decimated," which means that, as the rate becomes too high or the onboard solid-state recorder becomes too full, some of the recorded counts are removed. If decimation occurs, the fraction $(f_d - 1)/f_d$ of the counts below a decimation energy E_0 is removed. f_d is the decimation factor (weight), usually equal to 4 or 6. All events above E_0 are downlinked.[11] To prepare the synthetic count profiles, the number of counts in each bin was changed according to the Poisson distribution. The 1σ errors for non-decimated, fully decimated, and partially decimated data are $\sqrt{C}$, $\sqrt{f_d.C_{dc}}$, and $\sqrt{C_1 + f_d.C_{2,dc}}$, respectively. C is the measured count number in a bin for non-decimated data. C_{dc} is the count number in a bin of fully decimated data and consequently is corrected for this decimation. C_1 is the count number in the non-decimated portion and $C_{2,dc}$ is the corrected count number in the decimated portion of the measured rate in the case of partially decimated data. A detailed explanation is provided in the Appendix. The CCF was fitted for each of the 1001 synthetic profiles and for each burst in our sample. The median of such a distribution of 1001 maxima of polynomial fits was taken as the true lag L for each burst. These median lags L are used in the following statistical tests and are listed in Table 9. The 2.5% and 97.5% quantiles of such a distribution of 1001 maxima of polynomial fits for each GRB delimit the 95% CL statistical errors. These errors are also listed in Table 9.

We decided to calculate the spectral lags only for bursts with a signal-to-noise ratio higher than 3.5 in both channels. This signal-to-noise ratio is defined as $S_{T90}/\sqrt{S_{T90} + 2B_{T90}}$, where S_{T90} is a GRB signal over the background level B_{T90}, and both S and B are counts in a T_{90} time interval over the range 25 keV–1.5 MeV. The choice of this limit was made to ensure that the CCF was sufficiently smooth with a clear peak, allowing determination of a reliable lag. Therefore, excluding the noisiest data, the number of GRBs with calculated lags is 142. Their distribution is presented in Figure 3.

The GRB peak-count number S was derived from the light curve with the maximal count number C at the range 25 keV–1.5 MeV after subtracting the background B. The peak-count rate F is given as the peak-count number S divided by the width of the time bin δt_{res}. This width was different for different GRBs, and covered a range between 2 ms and 3 s. The dimensions of the peak-count rate are counts s^{-1}.

The 1σ error σ_F of the peak-count rate F was calculated as $\sigma_F = \sigma_S/\delta t_{res}$, where the error σ_S of the GRB peak-count number is $\sigma_S = \sqrt{(\sigma_C)^2 + (\sigma_B)^2}$. We assume that errors of the maximal count numbers σ_C and of the background $\sigma_B = \sqrt{B}$ are Poissonian and independent. The error σ_C is $\sigma_C = \sqrt{C}$ in the case of non-decimated data, given by expression (A2) in case of fully decimated data, and given by expression (A4) in case of partially decimated data (see the Appendix). The peak counts with errors were calculated for all 427 objects.

[11] http://sprg.ssl.berkeley.edu/~dsmith/hessi/decimationrecord.html

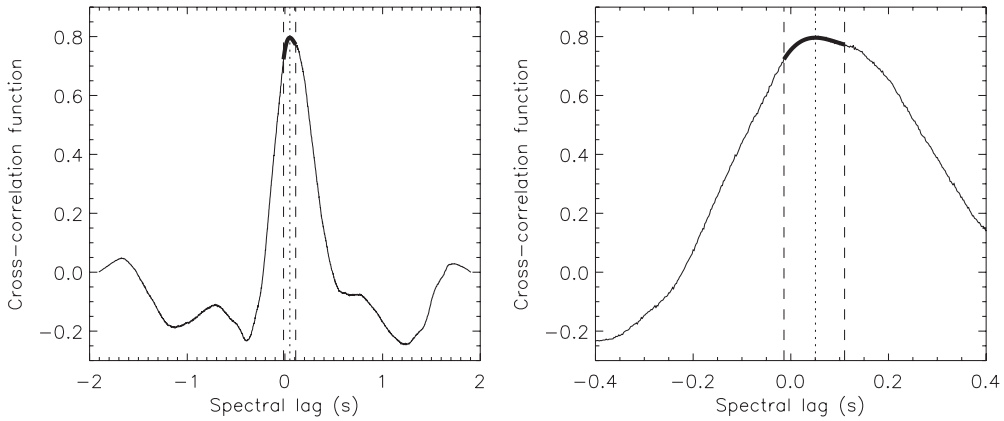

Figure 2. Left: an example of the cross-correlation function of two background-subtracted count light curves of the very bright GRB 060306 derived at two energy bands, 400–1500 keV and 25–120 keV. Right: detail from the same curve with the third-order polynomial fit (thick solid curve). The position of the maximum of the fit measures the spectral lag (dotted line). The boundaries of the polynomial fit are marked with dashed lines.

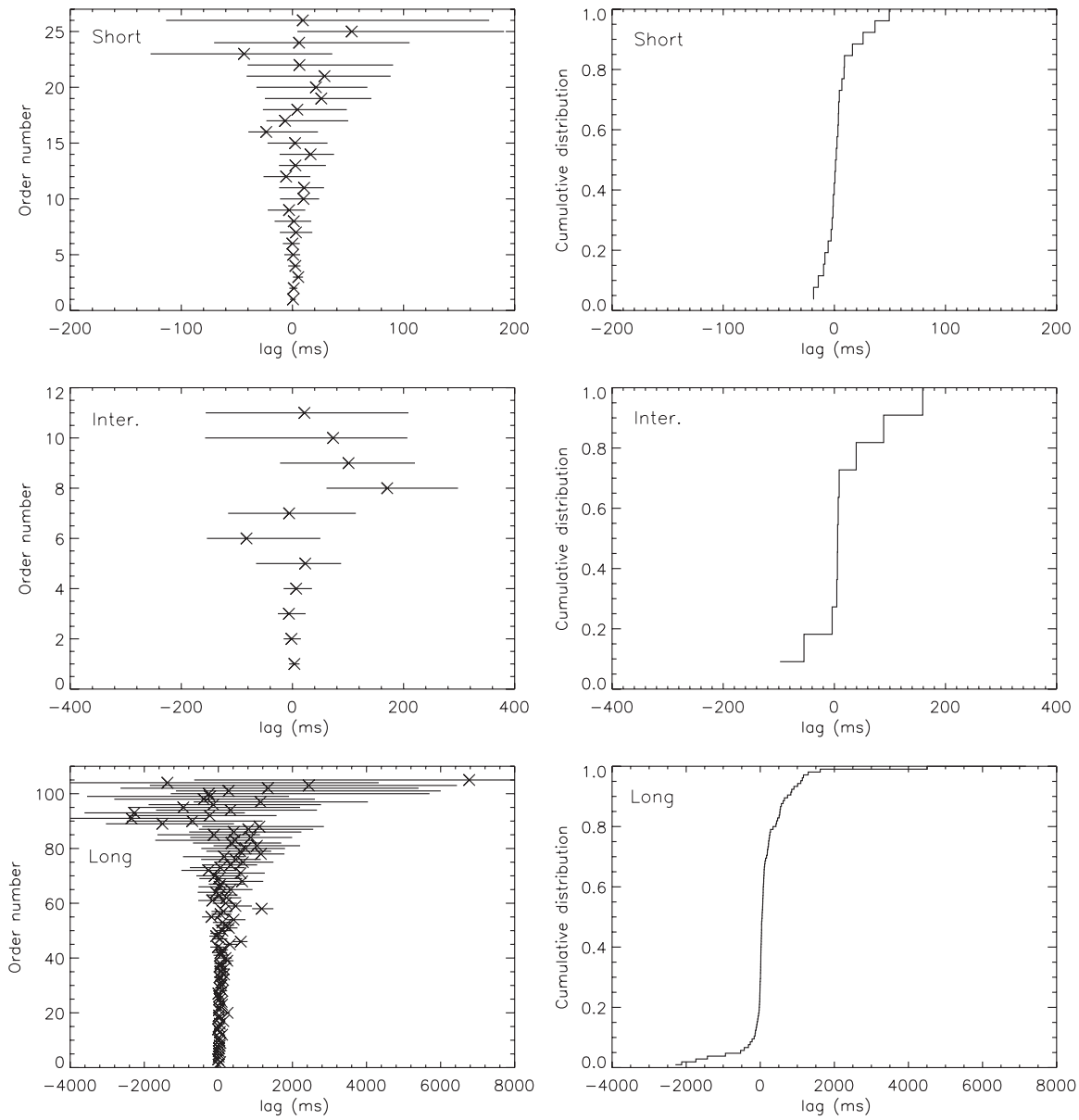

Figure 3. Left panels: the spectral lags of *RHESSI* GRBs sorted along the *y*-axis with respect to the value of (+error + |−error|) for short-, intermediate-, and long-duration bursts. The median lags for each GRB were taken from the lags of 1001 synthetic background-subtracted count time profiles obtained by Monte Carlo simulations of the measured profiles that were randomly influenced by the Poissonian noise. The error bars are composed of the 95% CL statistical error and the profile time resolution. A positive lag means that the low-energy counts are delayed. Right panels: the cumulative distributions of the median lags obtained for the three groups of bursts are shown.

3. PROPERTIES OF THE GRB GROUPS

3.1. Distribution of Spectral Lags

In this section we use the A-D test (Anderson & Darling 1952; Darling 1957) to compare distributions of spectral lags of different GRB groups (see Figure 3) found by the ML method applied to durations and hardness ratios (see Section 2). The short (intermediate, long) group contains 26 (11, 105) objects. The mean values of the spectral lags of these groups are similar; hence, we use the A-D test because it is particularly sensitive to the tails of the distributions tested (Scholz & Stephens 1987). For its calculation, we employ the *adk* package of the R software[12] (R Development Core Team 2011). The results are summarized in Table 2.

[12] http://cran.r-project.org

Table 2
A-D Tests of the Spectral Lag Distributions for the *RHESSI* Database

Groups	A-D P (%)	Group	Mean L (ms)	Median L (ms)	σ (ms)
Inter.–short	16.8	Short	4.9	1.9	16.7
Inter.–long	4.2	Inter.	28.7	5.9	78.4
Short–long	$<10^{-3}$	Long	178.0	50.8	874.9

Notes. Left: the results of the A-D tests are presented. The null hypothesis is that the two samples are drawn from the same distribution. P denotes the P-value of the test. Right: the means, medians, and standard deviations σ of the lags are listed.

The A-D test gives a significance of 16.8% (the probability that the two samples are drawn from the same distribution) for the short–intermediate pair, and it yields a significance of

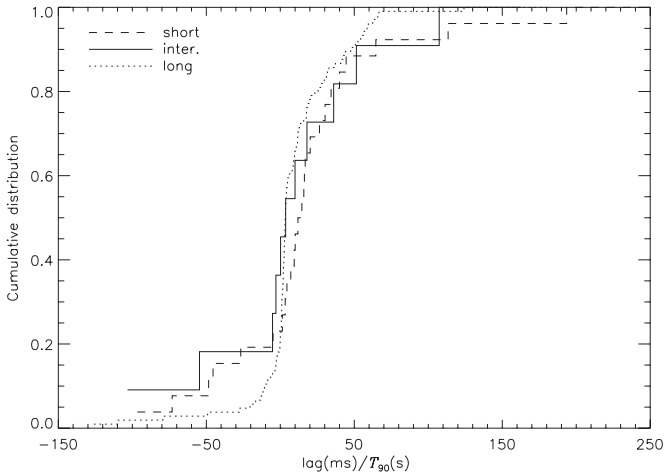

Figure 4. Cumulative distributions of the normalized lags for the three *RHESSI* GRB groups.

Table 3
A-D Tests of the Equality of the Normalized Lag Distributions Between Different *RHESSI* GRB Groups

Groups	A-D P (%)	Group	Mean L(ms)/T_{90}(s)	Median L(ms)/T_{90}(s)	σ
Inter.–short	54.2	Short	21.3	15.8	63.3
Inter.–long	45.0	Inter.	17.6	5.5	63.0
Short–long	6.0	Long	10.2	3.4	32.0

Notes. Left: the results of the A-D tests are presented. P denotes the P-value of the test. Right: the means, medians, and standard deviations σ of the normalized lags are also mentioned.

4.2% for the long–intermediate pair. Therefore, in the cases of short and intermediate groups, we cannot reject the null hypothesis that the two samples are drawn from the same distribution on a sufficiently low level (5%). On the other hand, this null hypothesis can be rejected in the case of long and intermediate groups, but the significance is not far below the 5% level. The same test applied to the lags of the short–long pair yields a significance of $<10^{-3}$%. Therefore, in this case, the null hypothesis can be rejected with a high significance. This strongly supports the well-known claim that the short and long GRBs are really different phenomena and confirms the results of Norris et al. (2001) (obtained with BATSE), but now by using the *RHESSI* instrument.

3.2. Distribution of Normalized Lags

In this section we compare the distributions of normalized lags (Figure 4), i.e., L/T_{90}, with the absolute values of the lags. Again, we use the A-D test between the different GRB groups mentioned in the previous section. The number of events within the groups is therefore the same. The results are summarized in Table 3.

The A-D test gives a significance level of 54.2% for the short–intermediate pair, and it gives a significance level of 45.0% for the long–intermediate pair. The significances are considerably above the 5% level; therefore, the null hypothesis that the samples are drawn from the same distribution cannot be rejected. For the short–long pair, the A-D test gives a significance level of 6.0%. Where the normalized lags are concerned, the difference between the short and long bursts is not definite.

Table 4
K-S Tests Applied to the Peak-count Rates F for the *RHESSI* Database

Groups	D	K-S P (%)	Group	Mean F (s^{-1})	Median F (s^{-1})	σ (s^{-1})
Inter.–short	0.44	0.9	Short	9 485	5 163	20 418
Inter.–long	0.55	3×10^{-5}	Inter.	4 412	2 546	5 586
Short–long	0.69	$<10^{-6}$	Long	2 589	1 038	7 673

Notes. Left: the results of the K-S tests are presented. The K-S distance D and the K-S significance P are given. Right: the means, medians, and standard deviations of the peak-count rates are listed.

Table 5
Monte Carlo Double-check of Results from the Statistical Tests

Tests	Inter.–Short	Inter.–Long	Short–Long
Lags	8 556 (85.6%)	6 938 (69.4%)	0 (0.0%)
Norm. lags	9 936 (99.4%)	8 862 (88.6%)	1 458 (14.6%)
Peak rates	47 (0.5%)	0 (0.0%)	0 (0.0%)

Notes. The number of cases out of 10,000 MC cycles (and their percentages) is given for A-D (lags and norm. lags) and K-S (peak rates) probability values exceeding 5% for tests done on spectral lags, normalized lags, peak-count rates, and different pairs of GRB groups.

3.3. Distribution of Peak Counts

Here, we used K-S test (Kolmogorov 1933; Smirnov 1948) to compare the cumulative distributions of the peak counts among the different GRB groups. The short (intermediate, long) group contains 42 (18, 367) objects. The results are presented in Table 4 and shown in Figure 5.

The results of the K-S tests imply that the distributions of the peak-count rates are different over all three groups. In particular, the K-S significance level for the intermediate versus short bursts is 0.9%, for intermediate versus long bursts it is 3×10^{-5}%, and for short versus long bursts it is $<10^{-6}$%.

3.4. Monte Carlo Simulations

In order to test the robustness of the results obtained by the A-D tests applied on lags L, normalized lags L/T_{90}, and K-S tests applied on peak-count rates F, one can use the MC method.

In the case of spectral lags we proceeded in the following way. The procedure described in Section 2—calculation of statistical errors of the lags by application of Poisson noise—provided a distribution of 1,001 lags for each GRB. Thus, for each GRB we randomly selected one lag from its distribution and made 10,000 data samples. Then, the A-D tests for these 10,000 samples were calculated.

In the case of peak rates, we proceeded as follows. We applied the Poisson noise to the measured light curves and subtracted the background in order to obtain the simulated data. Then we derived the peak-count rate for the same peak time when the peak was found in the measured light curves. We proceed in this way for each GRB. Afterward, we calculated K-S tests and repeated this sequence 10,000 times.

The number of cases in which the A-D and K-S probability reached values higher than 5% for tests done on different pairs of GRB groups is given in Table 5. The results of MC simulations comparing spectral lags and normalized lags are shown in Figure 6.

The MC method confirms that the distributions of spectral lags between short and long GRB groups are different. Let us compare results from the MC simulations of lags and normalized

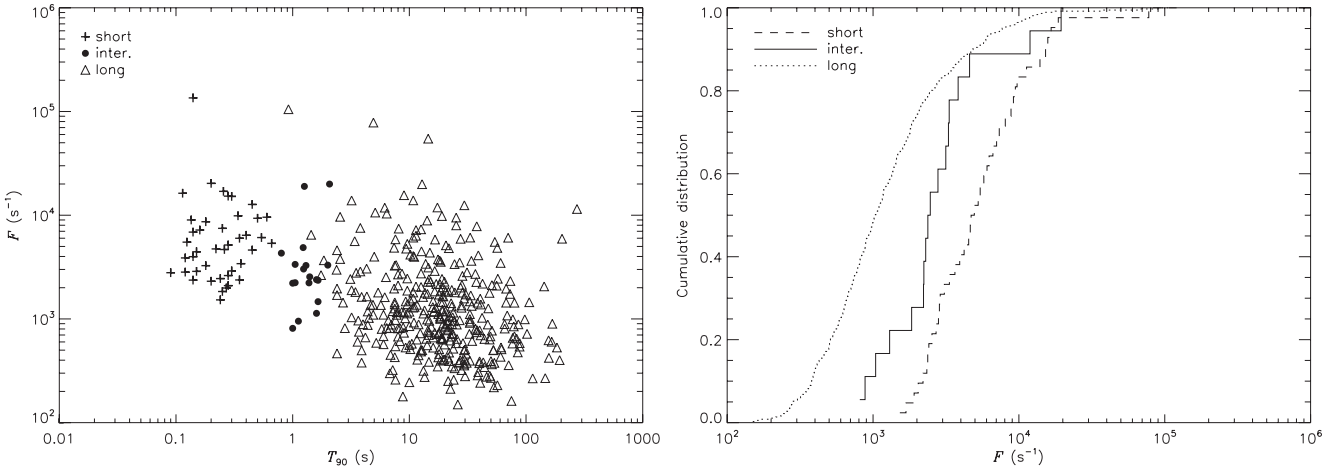

Figure 5. Left panel: peak-count rates F of *RHESSI* GRBs as a function of T_{90} durations for the three GRB groups, identified by the analysis of the hardnesses and durations, are displayed. Right panel: cumulative distributions of these peak-count rates F for the short-, intermediate-, and long-duration bursts are shown.

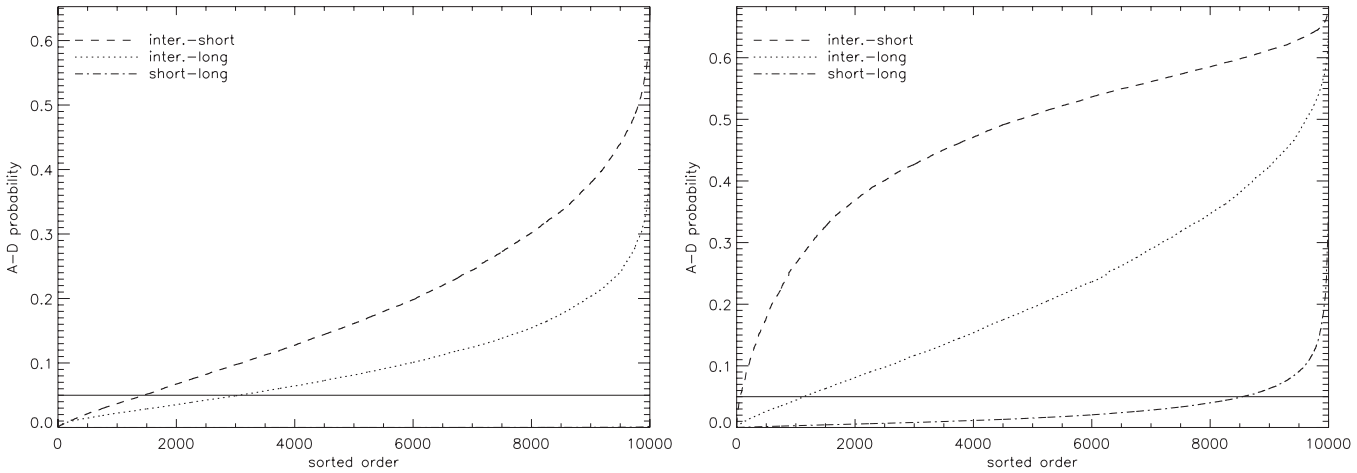

Figure 6. Left panel: The A-D probabilities of the tests applied to the samples of lags (left panel) and normalized lags (right panel) obtained from 10,000 MC cycles for different GRB groups. The horizontal solid line denotes the 5% threshold.

lags between the intermediate–short and intermediate–long pairs with the results of the tests applied directly to median lags (Table 2) and median normalized lags (Table 3). Then, one can see that MC simulations give an A-D probability >5% more often than expected. This can be caused by the fact that for some GRBs, weak and noisy ones, the distribution of lags found by the MC method might not follow the real distribution, because after applying the Poisson noise the polynomial fit of the CCF may not well describe the CCF peak. The reason for this conjecture is that the fitting range remained fixed and the same for the simulated data as for the measured data. In other words, the fitting range suitable for the measured data need not be suitable for the simulated data. In this case, we think that the A-D tests applied to the median lags give more reliable results than do the MC simulations. However, one mutual behavior is seen here: the intermediate–short pair has distributions of lags and normalized lags more similar than does the intermediate–long pair. This feature is seen both in the A-D tests applied to the median lags/normalized lags and in the A-D tests of the MC data samples.

MC simulations also confirm results of K-S tests applied directly to the measured peak rates. We can conclude that the short-, intermediate-, and long-duration bursts have different distributions of peak-count rates. The results of MC simulations comparing spectral lags and normalized lags are shown in Figure 7.

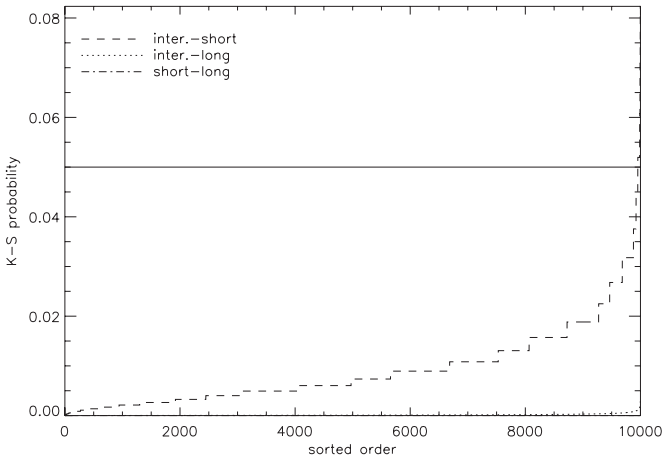

Figure 7. K-S probabilities of the tests applied to the samples of peak-count rates obtained from 10,000 MC cycles for different GRB groups.

4. DISCUSSION

4.1. Comparison with the BATSE Database

The lags of GRBs from the BATSE data set are different for the short and long groups (Norris et al. 2001): for the short bursts the lags on average are close to zero, but for the long bursts they are positive. Norris and his collaborators did not

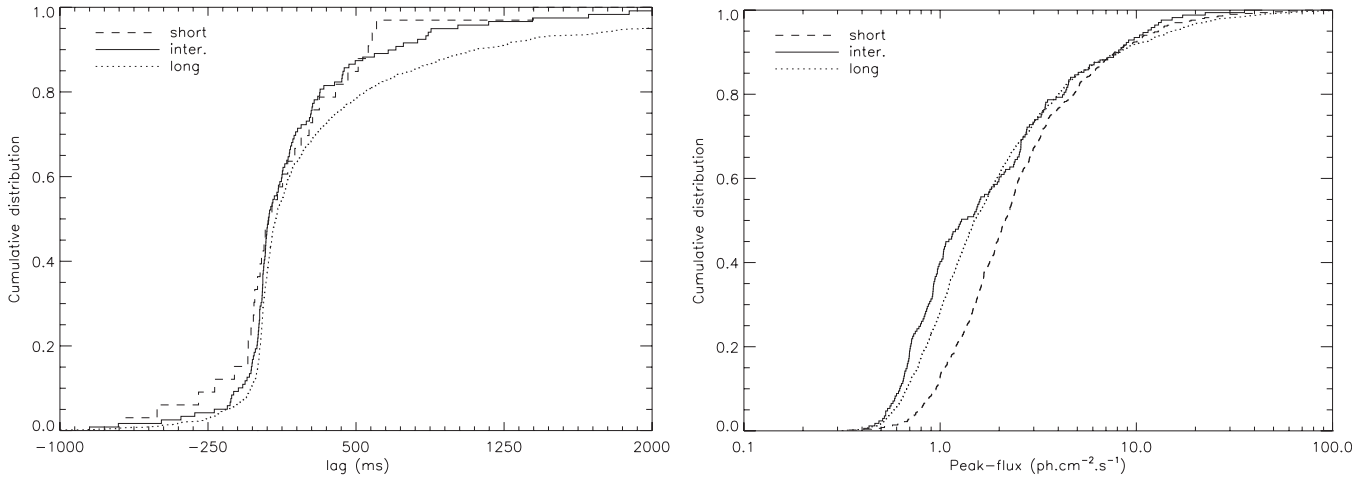

Figure 8. Cumulative distributions of the spectral lags (left panel) and peak fluxes (right panel) for the three BATSE GRB groups are shown.

Table 6
A-D Tests of the Equality of the Spectral Lag Distributions for the BATSE GRBs

Groups	A-D P (%)	Group	Mean L (ms)	Median L (ms)	σ (ms)
Inter.–short	51.3	Short	177.1	72.0	454.6
Inter.–long	3.8	Inter.	207.5	60.0	464.2
Short–long	9.7	Long	390.7	94.0	848.2

Notes. Left: the results of the A-D tests are presented. P denotes the P-value of the test. Right: the mean, median, and standard deviations σ of the lags.

Table 7
K-S Tests Applied to the Peak Fluxes of the BATSE GRBs

Groups	D	K-S P (%)	Group	Mean F	Median F	σ
Inter.–short	0.30	$<10^{-6}$	Short	4.00	2.15	6.80
Inter.–long	0.13	1.0	Inter.	3.15	1.29	4.91
Short–long	0.21	$<10^{-6}$	Long	4.09	1.51	10.31

Notes. Left: the results of the K-S tests applied to the peak fluxes F (photons cm^{-2} s^{-1}) are presented. The notation is as defined in Table 4. Right: the means, medians, and standard deviations of the peak fluxes are listed.

study the lags of the intermediate bursts separately (Norris et al. 2001; Norris 2002; Norris & Bonnell 2006). For the sake of completeness we have attempted to do this for the publicly available data. Horváth et al. (2006) define membership within the groups for all BATSE GRBs. Additionally, Norris (2002) defines the lag for any GRB with $T_{90} > 2$ s.[13] Compilation of these two lists and the application of the A-D test to the lags of the three BATSE groups (here for the first time) produced the results collected in Table 6 and shown in Figure 8. The short (intermediate, long) group contains 33 (119, 1179) objects here. Of course, one must keep in mind that this sample is drastically truncated for the short bursts. Hence, the short–intermediate and the short–long comparisons can serve only as qualitative indicators. Even the intermediate–long pair cannot be taken as representative because the truncation $T_{90} > 2$ s can also omit several intermediate GRBs.

Keeping all this in mind, if the lags are taken into consideration we can say that there is some similarity between the BATSE and the *RHESSI* databases. First, there is a similarity with regard to the intermediate–long pair: the difference is confirmed, though not at a high significance level; remarkably, the significances from the A-D test are comparable (3.8% and 4.2%). Second, there is a similarity with regard to the intermediate–short pair: in both databases the A-D test reveals that for these two groups the distributions of GRB lags are similar; the significances are 51.3% for BATSE and 16.8% for *RHESSI*. However, one must again keep in mind that our BATSE sample of short bursts is truncated, as is the sample of intermediate bursts. Third, both databases show a difference between the average lags for the short–long pairs (for the BATSE databases the difference between the distributions is shown to be insignificant; the A-D P-value is only 9.7%, probably as a result of the sample truncation, but Norris et al. (2001) make this claim unambiguously).

The results of the K-S tests applied to the peak fluxes of the 64 ms resolution light curves for the BATSE data imply that the distributions are different over all three groups. The short (intermediate, long) group contains 502 (169, 1282) objects here, and the K-S tests are summed in Table 7 and shown in Figure 8. These results are to be expected because, for example, Nakar (2007) claimed that the peak fluxes of short GRBs are roughly 20× smaller than those of the long ones. It is also known that the intermediate BATSE group is "intermediate" concerning the fluence (Mukherjee et al. 1998).

Therefore, our comparison of these *RHESSI* and BATSE groups finds similarities. In the case of the BATSE database, all three groups are different in respect to two quantities (duration and peak flux). It is remarkable that for BATSE the hardness of the intermediate group is strongly anticorrelated with the duration (Horváth et al. 2006). Since the hardness of the intermediate group differs from the hardnesses of the short and long ones, these studies support the opinion that all three BATSE groups represent different phenomena.

4.2. Comparison with the Swift Database

The lags of the GRBs from the *Swift* data set are also different for the short and long groups (de Ugarte et al. 2011). de Ugarte et al. (2011) also discussed the lags of the intermediate bursts, and they found a behavior which does not resemble the cases found in the *RHESSI* and BATSE data sets. *Swift*'s intermediate–long pair has on average similar lags, but there is a statistically significant difference in the short–intermediate pair. Thus, if the lags are considered, *Swift*'s intermediate group is similar to its long group (de Ugarte et al. 2011).

[13] http://heasarc.gsfc.nasa.gov/docs/cgro/analysis/lags/web_lags.html

On the other hand, the peak fluxes differ significantly in the short–intermediate and intermediate–long pairs, respectively. The peak fluxes of the short–long pair are not different from the statistical point of view (Veres et al. 2010). Nevertheless, de Ugarte et al. (2011) concluded that "*Swift*'s intermediate bursts differ from short bursts, but exhibit no significant differences from long bursts apart from their lower brightness." In other words, in the *Swift* database there is a clear similarity between the intermediate group and the long one. The physical difference between the short and long bursts in the *Swift* database also holds (Veres et al. 2010; de Ugarte et al. 2011).

Comparison with *Swift*'s groups leads to the conclusion that the third group in the *Swift* database is strongly related to the long group, as stated by Veres et al. (2010) and de Ugarte et al. (2011), and only the short group might represent another phenomenon. There is a difference in the hardness, peak flux, and duration for the intermediate–long pair (Horváth et al. 2008; Veres et al. 2010), but no clear separation occurs for the lags (de Ugarte et al. 2011). We have no reason to query the conclusions of Veres et al. (2010) that the intermediate group is related to XRFs, which in turn can be related to standard long GRBs. We add that the separation within the long group itself into harder and softer parts is not fully new (Pendleton et al. 1997; Tavani 1998). We allow the claim that the intermediate-duration bursts in the *RHESSI* and *Swift* databases are different phenomena. These results follow exclusively from the statistical analyses.

4.3. Discussion of the Number of Groups

In order to provide an extended discussion of the number of GRB groups, we apply clustering methods to our data sample. This also serves to extend the statistical analysis performed by Řípa et al. (2009). In general, the clustering methods can be divided into parametric and non-parametric types. Parametric methods assume that the data follow a pre-defined model (in our case a sum of multivariate Gaussian functions). These methods assign for each GRB a probability of membership in a certain group. The non-parametric methods, e.g., K-means clustering, provide definite assignments of each burst to a given group. More details about these methods can be found in the book by Everitt et al. (2011). Model-based clustering is also described in McLachlan & Peel (2000).

We apply model-based clustering and K-means clustering methods to our *RHESSI* data sample by using the algorithms implemented in the R software.

4.3.1. Model-based Clustering Method

In this method, we assume that the distributions of the parameters tested (logarithms of durations, hardness ratios, peak-count rates, and normalized lags) follow a superposition of Gaussian functions. A similar analysis for GRB classification was performed by Mukherjee et al. (1998), Horváth et al. (2006), and Veres et al. (2010).

The ML method is used to find the best-fit model parameters. Adding more free parameters to a fitted model can increase the likelihood, but may also result in overfitting. It is possible to penalize a model for more free parameters. This can be done by a method called the Bayesian information criterion (BIC), presented by Schwarz (1978). The function, which must be maximized to get the best-fit model parameters, is $\mathrm{BIC} = 2 \ln l_{\max} - m \ln N$, where $l_{\max}$ is the ML of the model, m is the number of free parameters, and N is the size of the sample. In our work we use the BIC to determine the most probable model, its parameters, and the number of its components.

For model-based clustering, we use the *Mclust* package[14] (Fraley & Raftery 2000) of R. For an explanation of the different models, see the Mclust manual.[15] The nomenclature of the different models in Mclust involves the following designations: the volumes, the shapes, and the orientation of the axes of all clusters may be equivalent (E) or may vary (V), and the axes of all clusters may be restricted to parallel orientations with the coordinate axes (I).

4.3.2. Model-based Clustering—Two Variables

First, we start with a two-dimensional case and fit T_{90} durations and hardnesses H. The data sample consists of 427 bursts (Table 7 of Řípa et al. 2009 and Table 1).

In this case, the number of free parameters of the model with k bivariate Gaussian components is $6k - 1$ ($2k$ means, $2k$ standard deviations, k correlation coefficients, and $k - 1$ weights, because the sum of the weights is 1). For the most general model, all parameters are free. However, sometimes we want to test models in which some of the parameters between different components are related to other parameters, e.g., all components have the same weight or shape, etc. In this case, the number of degrees of freedom is reduced.

As seen in Figure 9, the best-fit model has $k = 2$ components with equal volumes, variable shapes, and with the axes of all clusters parallel to the coordinate axes (EVI model). This best-fit model has a value of BIC $= -681.5$. The EVI model with one component gives BIC $= -899.1$, and the one with three components gives BIC $= -701.8$. For all other models tested with $k = 1$ component, the highest BIC is -820.3 and with $k = 3$ components -694.3, which are clearly below the maximum.

The difference between the BIC of two models gives us information about the goodness of fit. According to Kass & Raftery (1995) and Mukherjee et al. (1998), a difference in BIC of less than 2 represents weak evidence, a difference between 2 and 6 represents positive evidence, between 6 and 10 strong evidence, and a difference greater than 10 represents very strong evidence in favor of the model with the higher BIC.

In our case, the difference between the best-fit model (EVI) with two components and the EVI models with one or three components is always higher than 10. This gives strong support for the EVI model with $k = 2$ components.

The two components are the short/hard and long/soft groups. The intermediate-duration bursts shown in Figure 1 are assigned to the short/hard group by this test.

4.3.3. Model-based Clustering—Three Variables

Next, we perform model-based clustering of three variables: T_{90} durations, hardnesses H, and peak-count rates F. Since the peak rates were measured for all events, the sample here also consists of all 427 bursts (Table 9).

The best-fit model has $k = 3$ components (see Figure 9) with equal volumes, equal shapes, and equal correlation coefficients between all clusters (EEE model). This best model has a value of BIC $= -1156.6$. The EEE model with two components gives BIC $= -1168.7$, and for four components BIC $= -1174.6$. Markedly high values of BIC are also obtained for the EEI,

[14] http://cran.r-project.org/web/packages/mclust/index.html
[15] http://www.stat.washington.edu/research/reports/2006/tr504.pdf

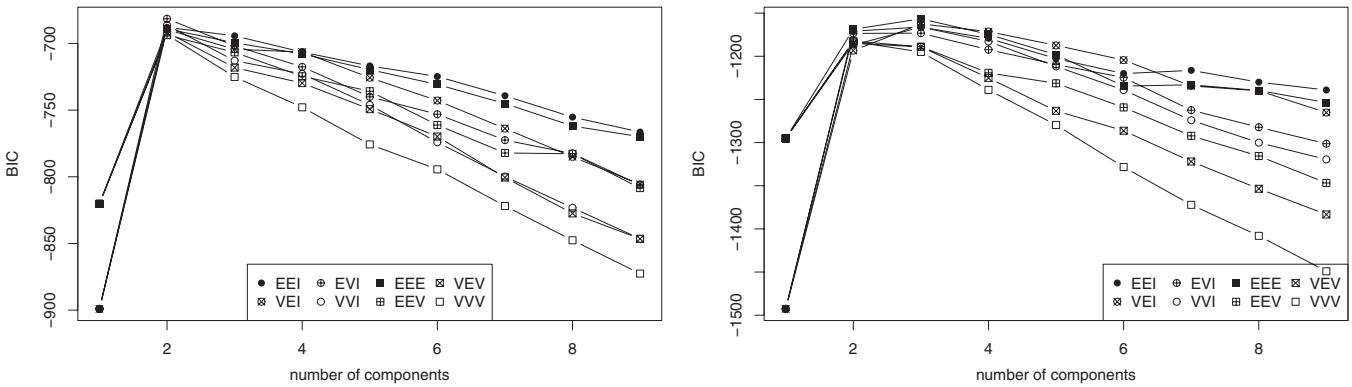

Figure 9. Left panel: Bayesian information criterion (BIC) values for different models as a function of the number of bivariate Gaussian components. The higher the BIC value, the more probable the model. The most probable model is EVI with two components. The data sample consists of two variables: T_{90} durations and hardness ratios. Right panel: BIC values for different models plotted against the number of components. The most probable model is EEE with three components. The data sample consists of three variables: T_{90} durations, hardness ratios H, and peak-count rates F.

VEI, and VVI models with $k = 3$ components: BIC $= -1166.2$, BIC $= -1162.5$, and BIC $= -1166.1$, respectively.

The difference in BIC between the EEE model with three components and the EEE models with two or four components is >10. The other models with other numbers of components (except the above-mentioned EEI, VEI, and VVI models with three components) give BIC values lower by at least 10. This provides strong evidence in favor of the EEE model with $k = 3$ components. The group structure of this model with three components is shown in Figure 10.

The intermediate-duration bursts shown in Figure 1 are assigned to the short/hard group by this test. A new result here is that the group of long bursts is separated into high and low peak-flux clusters.

4.3.4. Model-based Clustering—Four Variables

Spectral lags of BATSE GRBs, i.e., the time delay between low- and high-energy photons from short and long groups, have been found to differ. For short bursts, an average lag is ∼20–40 times shorter than for long bursts, and the lag distribution is close to symmetric about zero—unlike long bursts (Norris et al. 2001; Norris 2002; Norris & Bonnell 2006). This result gave us the idea to incorporate the spectral lags as well.

In this section, we apply the model-based clustering to GRB peak-count rates F, T_{90} durations, hardness ratios H, and a new addition to the variables, normalized lags L/T_{90}. Since the *RHESSI* spectral lags were calculated for only 142 bursts (Table 9), our sample is truncated.

The best-fit model has $k = 2$ components and is unconstrained, i.e., it has variable volumes, variable shapes, and variable correlation coefficients (VVV). The best BIC value for this model is BIC $= -1768.4$. However, the VVV model with $k = 3$ components gives a similar value of BIC $= -1768.5$. The other models give BIC values lower by at least 10. This strongly supports the VVV model with $k = 2$ components. There is no need to introduce the VVV model with three components, which has more free parameters. The two components are separated according to the values of normalized lags into zero- and non-zero-lag events.

4.3.5. Summary of Model-based Clustering

The model-based clustering of two-parameter data (T_{90} and H) gives strong evidence in favor of the EVI model with two components. The analysis of three-parameter data (T_{90}, H and F) shows that the best-fit model is EEE with three components. Surprisingly, a new result is obtained here: the group of long bursts is separated into high and low peak-flux clusters. The analysis of four-parameter data (T_{90}, H, F, and L/T_{90}) supports the VVV model with two components only. The separation into the two components here is according to the values of normalized lags into zero- and non-zero-lag events. The summary of the results from the model-based clustering is presented in Table 8.

Table 8
A Summary of the Results from the Model-based Clustering

	Model	k	BIC	ΔBIC $k=1$	ΔBIC $k=2$	ΔBIC $k=3$	ΔBIC $k=4$	Evidence
Two par.	EVI	2	−681.5	>10	×	>10		Very strong
Three par.	EEE	3	−1156.6		>10	×	>10	Very strong
Four par.	VVV	2	−1768.4	>10	×	>10		Very strong

Notes. The results for model-based clustering applied to two, three, and four parameters are presented. The values of BIC for the best-fitted models with k components are listed, as are the differences from the models with other numbers of components.

4.3.6. K-means Clustering

One of the non-parametric clustering methods is K-means clustering (MacQueen 1967). Before we use our data for this method we scale them, i.e., we subtract the mean value and then divide them by the standard deviation. The reason for this procedure is that the clustering algorithm is sensitive to the distance scale of the variables. For more details about the application of the K-means method in a similar analysis of GRB data, see, e.g., Chattopadhyay et al. (2007) or Veres et al. (2010). For this clustering method we use the *kmeans* package implemented in the R software.

To use the K-means method, one must set the number of clusters beforehand. Then, the corresponding number of centers is found by minimizing the sum of squared distances from each burst to the center of the group to which they belong. There is no precise way to determine the best number of clusters with this method. However, it has been suggested that if one plots the within-group sum of squares (WSS) as a function of the number of clusters, then an "elbow" will indicate the best number (Hartigan 1975). This method does not provide any

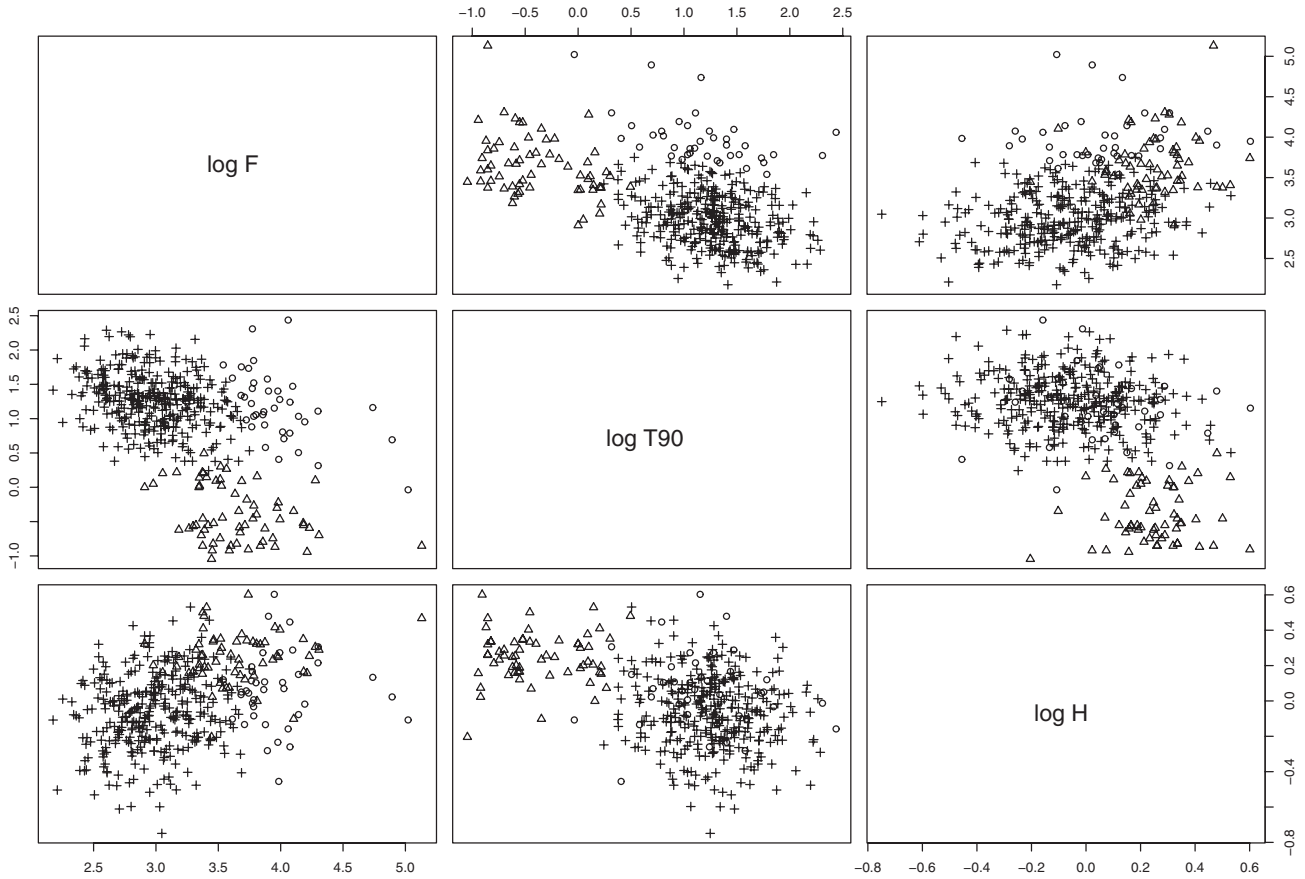

Figure 10. Scatter plot of 427 bursts, with measured T_{90}, H, and F assigned into three groups by the EEE model.

Table 9
The Spectral Lags and Peak-count Rates of the *RHESSI* GRBs

GRB[a]	Group[b]	L (ms)[c]	F (s^{-1})[d]	σ_F (s^{-1})[e]
020214	3	$42.4^{+56.7}_{-35.0}$	8885.9	221.3
020218	3	$607.0^{+181.9}_{-205.4}$	3630.7	92.9
020302	3		632.6	62.7
020306	1	$1.2^{+15.7}_{-17.3}$	9003.0	867.4
020311	3	$641.9^{+570.7}_{-519.2}$	1571.9	119.7
020313	3		891.2	89.3
020315	3		504.1	91.6
020331	3		307.2	71.8
020407	3		775.9	72.4
020409	3		268.6	56.0

Notes.
[a] *RHESSI* GRB number.
[b] The assignment to the GRB group—1: short, 2: intermediate, 3: long.
[c] Spectral lags were calculated from the difference in the count light curves at the energy intervals 400–1500 keV and 25–120 keV. The errors are the 95% CL statistical uncertainty and the light curve's time resolution.
[d] Peak-count rates derived in the 25–1500 keV band.
[e] 1σ statistical uncertainties of the peak-count rates.

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

probability indicating the significance or insignificance for the given best number of clusters.

The calculated WSS as a function of the number of groups for our data samples using two (T_{90}, H), three (T_{90}, H, F), and four (T_{90}, H, F, L/T_{90}) variables are rather smooth and do not demonstrate any remarkable and sharp "elbows;" thus, they do not provide useful information on the GRB classification.

4.4. Discussion of the Results

The K-S tests applied to peak-count rates show that the distributions are different over all three groups. The K-S significance level for the short–long pair is $<10^{-6}$%, for the intermediate–long pair is 3×10^{-5}%, and for the intermediate–short one is 0.9%. The short and long GRBs have clearly different distributions of peak rates. Also, the intermediate and long GRBs have clearly different distributions of peak rates. The intermediate–short pair also exhibits different distributions (K-S probability <5%), however less markedly than do the other pairs of groups. These results are confirmed by MC simulations.

The A-D tests applied to distributions of spectral lags unveil that the A-D probability for the short–long pair is $< 10^{-3}$%, for the intermediate–long pair the A-D probability is 4.2%, and for the intermediate–short one it is 16.8%. The short and long GRBs have clearly different distributions of spectral lags. The intermediate and long GRBs have A-D probability <5%; however, in this case the difference is not strong. The intermediate–short pair does not exhibit different distributions. The difference in the spectral lag distributions of the short–long pair of GRB groups is confirmed by MC simulations. In the cases of intermediate–short and intermediate–long pairs, the MC simulations reveal the same tendency as the A-D tests applied directly to the measured values, i.e., the intermediate–short pair has more similar distributions of spectral lags than does the intermediate–long pair. However, MC simulations give A-D

probability higher than 5% more often than expected. A possible reason is noted in Section 3.4.

The A-D tests applied to distributions of normalized lags show that these distributions cannot be claimed as different. The A-D probability for the short–long pair is 6.0%, for the intermediate–long pair is 45.0%, and for the intermediate–short one is 54.2%. Here, one can see the same tendency as in the case of A-D tests applied to spectral lags, i.e., the short–long pair has the least similar distributions of lags, the intermediate–long couple stays in the middle, and the intermediate–short pair has the most similar distributions. This tendency also appears in the MC simulations; however, the absolute frequency of the cases when A-D probability exceeds the 5% level happens more often than expected. The reason could be the same as in the case of MC simulations applied to absolute values of spectral lags.

The model-based clustering of two-parameter data (T_{90} and H) gives strong evidence in favor of a model with two components only. The two components are the short/hard and long/soft groups. The intermediate-duration bursts shown in Figure 1 are assigned to the short/hard group by this test. The analysis of three-parameter data (T_{90}, H, and F) shows that the best-fit model has three components. A surprising point here is that this method separates the group of long bursts into high and low peak-flux clusters. The analysis of four-parameter data (T_{90}, H, F, and L/T_{90}) supports a model with two components that are separated accordingly to the values of normalized lags into zero- and non-zero-lag events.

Surveying the A-D and K-S tests of the *RHESSI* data, it should be noted that the difference between the short and long bursts was again strongly confirmed. This follows from the different distributions of the spectral lags and from the different distributions of the peak-count rates; both results were confirmed by the MC method. This is already an expected result, but—usefully—in this case came from a new observational database.

According to Figure 1, the intermediate–short pairs of groups have similar hardness ratios. Also according to the results of the A-D test of the spectral lags, the distributions of lags are not different for the intermediate–short pair. However, the intermediate-duration and short-duration bursts are not completely the same because their peak-count rate distributions differ. On the other hand, the intermediate–long pairs of groups differ in hardness ratios, spectral lags, and peak-count rates. Therefore, in our opinion, it is possible that the intermediate group detected by *RHESSI* in Section 2 and by Řípa et al. (2009) may be a longer tail of standard short/hard bursts. This can also be supported by the fact that the model-based clustering method applied to hardness ratios and durations unveils only two clusters as the best solution: classical short/hard and long/soft groups; the intermediate-duration bursts are assigned to the short group.

The *RHESSI* intermediate and long groups seem to be different phenomena. This difference is supported by the distribution of the peak-count rates and spectral lags. The results show that the intermediate group is also "intermediate" with regard to its lags. The intermediate group detected by *Swift* was found to be related to XRFs (Veres et al. 2010), and those may in turn belong to the standard long GRBs (Kippen et al. 2003). In the case of *RHESSI*, the longer and softer GRBs are more difficult to detect, because *RHESSI*'s sensitivity declines rapidly below ≈50 keV and the weak and soft GRBs are not easily observable (Řípa et al. 2009). On the other hand, *Swift* is less sensitive in the photon energy range >150 keV. However, softer GRBs are readily detectable with this instrument. Hence, in our opinion, an instrumental effect may be responsible for the fact that the two satellites (*Swift* and *RHESSI*) detected different intermediate groups. This means that—from a statistical point of view—different groups can be found if one looks at different databases.

There are bursts observed with properties similar to the short bursts (hardness, lag), except their durations exceed 2 s. For example, in addition to Gehrels et al. (2006) and Kann et al. (2011) already mentioned in the Introduction, Norris & Bonnell (2006) claim that "short bursts with extended emission (SGRBEE) can have $T_{90} > 2$ s." Furthermore, others (de Barros et al. 2011) also propose the astrophysical fragmentation of the short GRB group.

Concerning SGRBEE, we inspected the light curves of all 18 *RHESSI* intermediate bursts but found no softer extended emission coming after the main hard spike as is typical for this kind of burst. Figure 3 of Perley et al. (2009) shows that the average T_{90} duration of the initial spike of an SGRBEE lies between the average durations of short and long bursts. If *RHESSI* detects only the hard initial spike, and the softer extended emission is lost in the noise, then the intermediate group detected might be polluted by these objects. Therefore, we also checked the light curves of the *RHESSI* intermediate bursts as observed by Konus-*Wind* (Aptekar et al. 1995), because it also has good sensitivity below 50 keV (its range is 10–10,000 keV). It has an overlap with the following *RHESSI* intermediate bursts: GRB 020819A, GRB 030410, GRB 040329, GRB 050530, GRB 070802, GRB 070824, and GRB 080408. However, no extended emission was observed by Konus-*Wind* for these seven bursts. This observation indicates the *RHESSI* intermediate GRBs should not be dominantly polluted by SGRBEEs.

Furthermore, there are also additional indications that GRBs that do not belong to the long+XRF pair category may originate from a broad range of astrophysical phenomena. For example, Mukherjee et al. (1998) found four subclasses in the BATSE database from the year 1998, but the fourth group was populated by a single GRB. From a statistical point of view, such an object is an outlier of uncertain origin. Likewise, similar situations exist concerning the objects GRB 060614 (Gehrels et al. 2006) and GRB 110328A (Cummings et al. 2011). Any study of such a single unusual object is beyond the scope of this article, which provides only statistical analyses.

5. CONCLUSIONS

The main results of this study can be summarized as follows.

1. The ML test in the duration–hardness plane of 427 *RHESSI* GRBs, taken from Řípa et al. (2009) but now with six events corrected for decimation, again exhibits a statistically significant third, intermediate in duration, group. This completes the work of Řípa et al. (2009) using the durations and hardnesses only.
2. The spectral lags and peak-count rates have been calculated for GRBs observed by the *RHESSI* satellite for the first time. The spectral lags were obtained for 142 objects, and the peak counts were obtained for all 427 GRBs. Hence, we constructed a new observational database for this satellite. Then, the three GRB subgroups were analyzed statistically with respect to these new spectral lags and peak-count rates.

3. The difference between short and long groups has been confirmed. Usefully, this result came from a new observational database.
4. K-S and A-D tests applied to spectral lags and peak-count rates indicate that the intermediate group in the *RHESSI* database might be a longer tail of the short group or at least has some properties in common with this short group. Contrary to this, the intermediate and the long groups are different.
5. The group of *RHESSI* intermediate-duration GRBs is not dominantly populated by SGRBEEs.
6. The intermediate-duration bursts found in the *RHESSI* and *Swift* databases seem to be represented by different phenomena.

We thank Z. Bagoly, L. G. Balázs, I. Horváth, and P. Mészáros for useful discussions and comments on the manuscript. We also thank R. Aptekar and T. L. Cline for providing the Konus-*Wind* data. Thanks are also due for the valuable remarks of the anonymous referees. This study was supported by the OTKA grant K77795, by the Grant Agency of the Czech Republic grant No. P209/10/0734, by the Research Program MSM0021620860 of the Ministry of Education of the Czech Republic, and by Creative Research Initiatives (RCMST) of MEST/NRF and the World Class University grant No. R32-2008-000-101300.

APPENDIX

UNCERTAINTIES IN DECIMATED DATA

This Appendix describes the 1σ uncertainties in the bin counts for the decimated *RHESSI* data (Curtis et al. 2002; Smith et al. 2002).

A.1. Full Decimation

The calculation of 1σ uncertainty $\sigma_{C_{dc}}$ for the bin counts C_{dc} of the fully decimated data and then corrected for this decimation is as follows. For corrected bin counts C_{dc} it holds that

$$C_{dc} = f_d.C_d, \tag{A1}$$

where f_d is the decimation factor (weight), usually equal to 4 or 6 for the *RHESSI* data, and C_d is the number of counts in a bin of the decimated signal. If we assume that the counts in a bin follow Poisson statistics, then

$$\sigma_{C_{dc}} = \left|\frac{\partial C_{dc}}{\partial C_d}\right| \sigma_{C_d} = f_d\sqrt{C_d} = \sqrt{f_d.C_{dc}}, \tag{A2}$$

where σ_{C_d} is the dispersion of C_d.

A.2. Partial Decimation

Now consider the situation in which counts in a bin are only partially decimated, i.e., they consist of the non-decimated counts C_1 and the decimated signal $C_{2,d}$. This situation may happen when we sum counts over the energy band $[E_1; E_2]$, $E_1 < E_0 < E_2$, and only the counts below the energy E_0 are decimated. Then, the corrected signal C_{dc} is equal to

$$C_{dc} = C_1 + C_{2,dc} = C_1 + f_d.C_{2,d}, \tag{A3}$$

where f_d is again the decimation factor, and $C_{2,dc}$ is the corrected part of the signal that was decimated. The 1σ uncertainty $\sigma_{C_{dc}}$ is then given by

$$\sigma_{C_{dc}} = \sqrt{\left(\frac{\partial C}{\partial C_1}\right)^2 \sigma_{C_1}^2 + \left(\frac{\partial C}{\partial C_{2,d}}\right)^2 \sigma_{C_{2,d}}^2} = \sqrt{\sigma_{C_1}^2 + f_d^2.\sigma_{C_{2,d}}^2} = \sqrt{C_1 + f_d^2.C_{2,d}} = \sqrt{C_1 + f_d.C_{2,dc}}, \tag{A4}$$

where $\sigma_{C_1} = \sqrt{C_1}$ is the dispersion of the non-decimated part of the bin counts and $\sigma_{C_{2,d}} = \sqrt{C_{2,d}}$ is the dispersion of the decimated part of the bin counts.

REFERENCES

Anderson, T. W., & Darling, D. A. 1952, Ann. Math. Stat., 23, 193
Aptekar, R. L., Butterworth, P. S., Cline, T. L., et al. 1998, in AIP Conf. Proc. 428, Fourth Huntsville Gamma-Ray Burst Symposium, ed. C. A. Meegan, T. M. Koshut, & R. D. Preece (Melville, NY: AIP), 10
Aptekar, R. L., Frederiks, D. D., Golenetskii, S. V., et al. 1995, Space Sci. Rev., 71, 265
Bagoly, Z., Mészáros, A., Balázs, L. G., et al. 2006, A&A, 453, 797
Balastegui, A., Ruiz-Lapuente, P., & Canal, R. 2001, MNRAS, 328, 283
Balázs, L. G., Bagoly, Z., Horváth, I., Mészáros, A., & Mészáros, P. 2003, A&A, 401, 129
Balázs, L. G., Mészáros, A., & Horváth, I. 1998, A&A, 339, 1
Balázs, L. G., Mészáros, A., Horváth, I., & Vavrek, R. 1999, A&AS, 138, 417
Bellm, E. C., Bandstra, M. E., Boggs, S. E., et al. 2008, in AIP Conf. Proc. 1000, Gamma-Ray Bursts 2007: Proceedings of the Santa Fe Conference, ed. M. Galassi, D. Palmer, & E. Fenimore (Melville, NY: AIP), 154
Chattopadhyay, T., Misra, R., Chattopadhyay, A. K., & Naskar, M. 2007, ApJ, 667, 1017
Cummings, J. R. 2011, GCN, 11823, 1
Curtis, D. W., Berg, P., Gordon, D., et al. 2002, Sol. Phys., 210, 115
Darling, D. A. 1957, Ann. Math. Stat., 28, 823
de Barros, G., Amati, L., Bernardini, M. G., et al. 2011, A&A, 529, A130
de Ugarte Postigo, A., Horváth, I., Veres, P., et al. 2011, A&A, 525, A109
Everitt, B. S., Landau, S., Leese, M., & Stahl, D. 2011, Cluster Analysis (5th ed.; Wiley Series in Probability and Statistics; Chichester, UK: Wiley)
Foley, S., McGlynn, S., Hanlon, L., McBreen, S., & McBreen, B. 2008, A&A, 484, 143
Foley, S., McGlynn, S., Hanlon, L., et al. 2009, in Baltic Astronomy, Proceedings of the INTEGRAL/BART Workshop 2009, Vol. 18, ed. R. Hudec, J. Řípa, M. Kocka, & R. Havlíková (Vilnius, Lithuania: Inst. of Theoretical Physics and Astronomy, Vilnius Univ.), 279
Fox, D. B., Frail, D. A., Price, P. A., et al. 2005, Nature, 437, 845
Fraley, C., & Raftery, A. E. 2000, J. Am. Stat. Assoc., 97, 611
Gehrels, N., Norris, J. P., Barthelmy, S. D., et al. 2006, Nature, 444, 1044
Gendre, B., Galli, A., & Piro, L. 2007, A&A, 465, L13
Hakkila, J., Haglin, D. J., Pendleton, G. N., et al. 2000, ApJ, 538, 165
Hartigan, J. A. 1975, Clustering Algorithms (New York: Wiley)
Horváth, I. 1998, ApJ, 508, 757
Horváth, I. 2002, A&A, 392, 791
Horváth, I. 2009, Ap&SS, 323, 83
Horváth, I., Bagoly, Z., Balázs, L. G., et al. 2010, ApJ, 713, 552
Horváth, I., Balázs, L. G., Bagoly, Z., Ryde, F., & Mészáros, A. 2006, A&A, 447, 23
Horváth, I., Balázs, L. G., Bagoly, Z., & Veres, P. 2008, A&A, 489, L1
Huja, D., Mészáros, A., & Řípa, J. 2009, A&A, 504, 67
Kann, D. A., Klose, S., Zhang, B., et al. 2011, ApJ, 734, 96
Kass, R. E., & Raftery, A. E. 1995, J. Am. Stat. Assoc., 90, 773
Kippen, R. M., Woods, P. M., Heise, J., et al. 2003, in AIP Conf. Proc. 662, Gamma-Ray Burst and Afterglow Astronomy 2001: A Workshop Celebrating the First Year of the HETE Mission, ed. G. R. Ricker & R. K. Vanderspek (Melville, NY: AIP), 244
Kolmogorov, A. 1933, Giornale dell'Istituto Italiano degli Attuari, 4, 83
Kouveliotou, C., Meegan, C. A., Fishman, G. J., et al. 1993, ApJ, 413, L101
Lin, R. P., Dennis, B. R., Hurford, G. J., et al. 2002, Sol. Phys., 210, 3
Litvin, V. F., Matveev, S. A., Mamedov, S. V., & Orlov, V. V. 2001, Astron. Lett., 27, 416
MacQueen, J. 1967, in Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, ed. L. M. Le Cam & J. Neyman (Berkeley, CA: Univ. California Press), 281
Mazets, E. P., Golenetskii, S. V., Ilinskii, V. N., et al. 1981, Ap&SS, 80, 3

McLachlan, G., & Peel, D. 2000, Finite Mixture Models (Wiley Series in Probability and Statistics; New York, NY: Wiley-Interscience)
Mészáros, A., Bagoly, Z., & Vavrek, R. 2000, A&A, 354, 1
Mészáros, A., & Štoček, J. 2003, A&A, 403, 443
Mukherjee, S., Feigelson, E. D., Jogesh Babu, G., et al. 1998, ApJ, 508, 314
Nakar, E. 2007, Phys. Rep., 442, 166
Norris, J. P. 2002, ApJ, 579, 386
Norris, J. P., & Bonnell, J. T. 2006, ApJ, 643, 266
Norris, J. P., Cline, T. L., Desai, U. D., & Teegarden, B. J. 1984, Nature, 308, 434
Norris, J. P., Marani, G. F., & Bonnell, J. T. 2000, ApJ, 534, 248
Norris, J. P., Scargle, J. D., & Bonnell, J. T. 2001, in Gamma-Ray Bursts in the Afterglow Era: Proc. Int. Workshop held in Rome, 17–20 October 2000, ed. E. Costa, F. Frontera, & J. Hjorth (Heidelberg: Springer), 40
O'Shaughnessy, R., Belczynski, K., & Kalogera, V. 2008, ApJ, 675, 566
Pendleton, G. N., Paciesas, W. S., Briggs, M. S., et al. 1997, ApJ, 489, 175
Perley, D. A., Metzger, B. D., Granot, J., et al. 2009, ApJ, 696, 1871
Rajaniemi, H. J., & Mähönen, P. 2002, ApJ, 566, 202
R Development Core Team 2011, R: A Language and Environment for Statistical Computing (Vienna, Austria: R Foundation for Statistical Computing)
Řípa, J., Mészáros, A., Wigger, C., et al. 2009, A&A, 498, 399
Řípa, J., Veres, P., & Wigger, C. 2010, The Shocking Universe—Gamma Ray Bursts and High Energy Shock Phenomena, Proc. Conf. held in San Servolo, Venice, Italy, 14–18 September 2009, ed. G. Chincarini et al. (Bologna: Soc. Ital. Fisica), Vol. 102, 569
Sakamoto, T., Barthelmy, S. D., Barbier, L., et al. 2008, ApJS, 175, 179
Sakamoto, T., & Gehrels, N. 2009, in AIP Conf. Proc. 1133, Gamma-Ray Burst: Sixth Huntsville Symp., ed. C. Meegan, C. Kouveliotou, & N. Gehrels (Melville, NY: AIP), 112
Scholz, F. W., & Stephens, M. A. 1987, J. Am. Stat. Assoc., Vol. 82, Issue 399
Schwarz, G. 1978, Ann. Stat., 6, 461
Soderberg, A. M., Kulkarni, S. R., Nakar, E., et al. 2006, Nature, 442, 1014
Smirnov, N. V. 1948, Ann. Math. Stat., 19, 279
Smith, D. M., Lin, R. P., Turin, P., et al. 2002, Sol. Phys., 210, 33
Tavani, M. 1998, ApJ, 497, L21
Vavrek, R., Balázs, L. G., Mészáros, A., Horváth, I., & Bagoly, Z. 2008, MNRAS, 391, 1741
Veres, P., Bagoly, Z., Horváth, I., Mészáros, A., & Balázs, L. G. 2010, ApJ, 725, 1955
Veres, P., Řípa, J., & Wigger, C. 2009, eConf Proceedings C091122 of the 2009 Fermi Symposium, arXiv:0912.3919
Zhang, B. 2006, Nature, 444, 1010
Zhang, B., & Mészáros, P. 2002, ApJ, 581, 1236
Zhang, B., Zhang, B.-B., Virgili, F. J., et al. 2009, ApJ, 703, 1696