Astron. Astrophys. 345, 121–126 (1999)

ASTRONOMY AND ASTROPHYSICS

Bayesian flare event detection

ROSAT X-ray observations of the UV Cetus type star G 131−026

V. Hambaryan[⋆]**, R. Neuhäuser, and B. Stelzer**

Max-Planck-Institut für Extraterrestrische Physik, D-85740 Garching, Germany

Received 18 January 1999 / Accepted 10 February 1999

Abstract. We have analyzed *ROSAT* PSPC observations for flare detection using a formalism of Bayesian model comparison. The method is applicable to the direct unbinned data and divides the whole observation into the most probable partitioning of non-uniform segments during which the photon arrival rate had no statistically significant variation. If more than one such segment is found with significantly different count rates, then variability is detected. As an example for this method, we show how a flare on the X-ray source RX J0008.8+2050, the counterpart of the UV Cet type star G 131−026, located in the solar vicinity, is detected and various items relevant to flare parameter estimation are discussed.

Key words: methods: statistical – stars: flare – stars: individual: G 131-026 – stars: variables: general – X-rays: stars

1. Introduction

The flare phenomenon as a short term variability of brightness of stars is thought to be produced by rapid energy release and is observed in almost all types of stars. The presence of flares in low-mass stars is supported by many observational facts at many wavelengths and is well documented. Stellar flares were discovered at optical wavelengths in the 1940 s on the so called UV Cet type variables. The quiescent radiation of these stars is interrupted by randomly distributed flares. UV Cet type stars have very low luminosities and are seen in the vicinity of the Sun (Gershberg et al. 1998). Further, systematical investigations of stellar flares in nearby stellar clusters (e.g., Pleiades) and associations (e.g., Orion) have revealed a fundamental significance of this phenomenon for the physics and evolution of the young low-mass stars and protostars (see, e.g. Mirzoyan 1995, Montmerle 1997). At present, the observations of stellar flares at different wavelengths of the electromagnetic spectrum has become a scientifically productive area of astrophysics.

In the 1970 s, the first X-ray studies of stellar flares were made, in particular with the *Einstein Observatory* (*EO*) (1978-1981) and the *EXOSAT* Observatory (1983-1985). Comprehensive reviews of the *EO* observations of stellar X-ray flares are presented by Haisch (1983) and of the *EXOSAT* observations by Pallavicini et al. (1990). X-ray observations of stellar flares obtained during the *ROSAT* All-Sky Survey as well as in *ROSAT* pointings are discussed by Schmitt (1994).

However, to date no systematic investigation of flare stars and stellar flares based on *ROSAT* observations has been done. X-ray observations by *ROSAT* are very well suited for statistical examination: Several physical quantities of individual X-ray photons are registered, namely the time of photon arrival, the direction of incidence, and the photon energy.

On the other hand, the wide spread timing analysis methods of datasets based on time binning technique give rise to certain difficulties: First of all, one has no a priori knowledge of the relevant timescales of a given flare to be detected, and therefore many different binnings of the data have to be considered; second, the bins must be large enough so that there will be enough photons to provide a good statistical sample, while larger bins will dilute short flares; moreover, the common practice of binning data overlooks a considerable amount of information and introduces a dependency of results on the sizes and locations of the bins. Meanwhile, it has been known for some time (Cash 1979), that fitting models to unbinned data is quite straightforward (see, for discussion, Scargle & Bapu 1998).

Hence, it is worthwhile to examine the available observational data of *ROSAT* in a straightforward and systematic way for variability testing, in particular, for UV Cet type flare stars.

For the flare event detection, we shall use the method described by Scargle (1998) based on Bayesian statistics due to the nearly ideal Poisson nature of photon registration by the *ROSAT* detectors (Trümper 1982, Pfeffermann et al. 1986).

In this article we will try to address various items relevant to detections and studies of flares with *ROSAT*. First, we will describe the flare event detection algorithm applied to a typical *ROSAT* observational data set; second, we will demonstrate this technique with an example, a flare event detected in *ROSAT* PSPC pointing observations on the nearby flare star No. 3 included in the Catalogue of UV Cet type stars and related objects (Gershberg et al. 1998).

Send offprint requests to: V. Hambaryan (vhambary@mpe.mpg.de)

[⋆] Visiting Astronomer from Byurakan Astrophysical Observatory, Armenia

2. Data reduction

2.1. Source detection

We investigated the *ROSAT* PSPC pointings in the direction of star No. 3 of the catalogue of UV Cet type stars and related objects of the solar vicinity (Gershberg et al. 1998), namely ROSAT pointings number 150026p and 700101p. First, the data were reduced in the manner similar to that described in detail by Neuhäuser et al. (1995). The X-ray sources for these two *ROSAT* PSPC pointings were identified using LDETECT, MDETECT and MAXLIK algorithms of EXSAS, the Extended Scientific Analysis Software System (Zimmerman et al. 1998) which runs under European Southern Observatory Munich Image Data Analysis System (ESO-MIDAS). In the LDETECT (local detect) algorithm a window is slid over the image. The source counts are determined within the window and background counts from an area surrounding it while the MDETECT (map detect) algorithm uses the background from the smoothed background map, which was created by masking the image with sources found in the previous step. As a result of these procedures the probability of a source existence is computed as a probability that the detected counts within the window and corresponding background do come from the Poisson distribution with the same expectation value whatever this expectation value is. Further, the merged source list derived from the above mentioned two algorithms is used in the maximum likelihood analysis (MAXLIK) for source detection, which works by varying parameters specifying source counts, extent and position until the likelihood is maximized. More detailed descriptions of these algorithms can be found in Cruddace et al. (1988) and Zimmerman et al. (1998). As a final result of this spatial analysis a list of X-ray sources having a maximum likelihood (ML) of existence larger than some threshold (the maximum likelihood can be converted into probability of existence through the equation $P = 1 - exp(-ML)$) is created.

2.2. Variability testing

As mentioned in the introduction, for flare event detection we have used a method developed by Scargle (1998) based on Bayesian statistics. The method is applicable to data that are known to be from a nearly ideal Poisson process, i.e. a class of independent, identically distributed processes, having zero lengths of dead time.

The data gathered in *ROSAT* observations (telescope+detector) allow the measurement of arrival times of individual X-ray photons to the nearest 1.22×10^{-4} s (i.e., $\delta t = \frac{1}{8192}$ s), an accuracy which is much smaller than the shortest time scales that are considered to be responsible for flare events on stars. Another departure from Poisson ideal distribution of counts is owing to the detectors finite dead time (see, for discussion, Scargle & Bapu 1998).

Nevertheless, the arrival times of photons registered by *ROSAT* and presented in so called Photon Events Tables in this context can be considered as close to a Poisson process, i.e. the arrival of a photon in any interval is independent of that in any other non-overlapping one.

Scargle's (1998) method is designed for application to photon counting data and uses it for decomposition into a piecewise constant Poisson process. For example, let us assume that during a continuous observational interval of length T, consisting of m discrete moments in time, at which it was possible to make measurements (spacecraft's "clock tick" $m = T/\delta t$), a set of photon arrival times $D\,(t_i, t_{i+1}, ..., t_{i+n})$ is registered. Suppose now that we want to use these data to compare two competing hypotheses, The first hypothesis is that the data are generated from constant rate Poisson process (model M_1) and the second one from two-rate Poisson process (model M_2). Evidently, model M_1 is described by parameter ν of one rate Poisson process while the model M_2 by parameters ν_1, ν_2 and τ describing two different parts of dataset D divided by any point τ from observational interval T (with lengths T_1 and T_2) at which the Poisson process switches from parameter (count rates) ν_1 to ν_2.

By taking as a background information (I) the proposition that one of the models under consideration is true and by using Bayes' theorem we can calculate the posterior probability (the probability that M_k ($k = 1, 2$) is the correct model) of each model by (see, e.g., Jaynes 1997)

$$P(M_k|D, I) = \frac{P(D|M_k, I)}{P(D|I)} P(M_k|I) \tag{1}$$

where $P(D|M_k, I)$ is the (marginal) probability of the data given M_k, and $P(M_k|I)$ is the prior probability of model M_K ($k = 1, 2$). The term in the denominator is a normalization constant, and we may eliminate it by calculating the ratio of the posterior probabilities instead of probabilities directly. Indeed, the extent to which the data support model M_2 over M_1 is measured by the ratio of their posterior probabilities and is called posterior odds ratio

$$O_{21} \equiv \frac{P(M_2|D, I)}{P(M_1|D, I)} = \left[\frac{P(D|M_2, I)}{P(D|M_1, I)}\right] \left[\frac{P(M_2|I)}{P(M_1|I)}\right]. \tag{2}$$

The first factor on the right-hand side of Eq. (2) is the ratio of the *integrated* or *global* likelihoods of the two models and is called the *Bayes factor* for M_2 against M_1, denoted by B_{21}. The global likelihood for each model can be evaluated by integrating over nuisance parameters and the final result for discrete Poisson events can be represented by (see, for details, Scargle 1998)

$$B_{21} = \frac{1}{B(n+1, m-n+1)} \sum B(n_1+1, m_1-n_1+1) \times B(n_2+1, m_2-n_2+1)\Delta\tau\,, \tag{3}$$

where B is the *beta function*, n_j and m_j, $(j = 1, 2)$, respectively are the number of recorded photons and the number of "clock tick" in the observational intervals of lengths T_1 and T_2. $\Delta\tau$ is the time interval between successive photons, and the sum is over the photons' index.

The second factor on the right-hand side of Eq. (2) is the prior odds ratio, which will often be equal to 1 (see below), representing the absence of a priori preference for either model.

It follows that the Bayes factor is equal to the posterior odds when the prior odds is equal to 1. When $B_{21} > 1$, the data favor M_2 over M_1, and when $B_{21} < 1$ the data favor M_1.

If we have calculated the odds ratio O_{21}, in favor of model M_2 over M_1, we can find probabilities for model M_2 by inverting Eq. (2), giving

$$P(M_2|D,I) = \frac{O_{21}}{1+O_{21}}\,. \quad (4)$$

Applying this approach iteratively to the observational data set, the Scargle (1998) method returns an array of rates, $(\lambda_1, \lambda_2, ..., \lambda_{cp})$, and a set of so called "change points" $(\tau_1, \tau_2, ..., \tau_{cp-1})$, giving the times when an abrupt change in the rate is determined, i.e. a significant variation. This is the most probable partitioning of the observational interval into segments during which the photon arrival rate was discernibly constant, i.e. had no statistically significant variations. Unlike most, this method does not stipulate or predetermine time bins – instead the data themselves determine an effective, non-uniform binning in time. Therefore this data analysis procedure does not itself impose a lower limit to the time scale on which variability can be detected. There are two free parameters in the method that are used to halt the segmentation process: The first is the minimum number of events that are allowed in a block (we have chosen two) and the second is a prior odds ratio that may be applied to disfavor segmentation. The prior odds ratio (second factor in the right-hand side of Eq. (2)) represents the relative likelihood assigned to the two models before the data is considered. Although this would appear to warrant a value of unity, in practice a larger value is used to prevent the method from making an incorrect decision to segment when two models are almost equally likely.

To have strong evidence in favor of segmenting, Scargle (1998) suggests to use as a prior odds ratio a quantity which is equal to the ratio of the length of the observational interval and the desired time resolution of the data. For interpreting *Bayes factors* (B_{21}) considered above, there are some rules of thumb for evidence of model M_2 against M_1 (Jeffreys 1961, Raftery 1994). Furthermore, it should be noted, that this point is not crucial – we find very similar results independent of the choice of the prior odds ratio.

As is well known, owing to the spacecraft orbits, *ROSAT* pointing observations contain data gaps. The integrated likelihoods that are used to decide between the above mentioned models M_1 and M_2 do not depend on the photon arrival times, only on the number of observing time resolution elements (number of *ROSAT* "clock tick"), and the number of registered photons in the observational interval (see, Eq. (3)). Therefore, in this context, the correct treatment of data gaps is to ensure that the gap length is not counted in the number of "clock tick" (numbers m_j in Eq. (3)). In the case of *ROSAT* pointed observations by PSPC or HRI, we allowed to take into account this point with a special descriptor available in the Photon Events Tables (see Zimmermann et al. 1998).

A number of X-ray sources, having maximum likelihood of existence $ML > 50$, were chosen for variability testing. For this purpose, we have extracted from the original observational data the part corresponding to a given source as follows: Around the center of each source a set of photon events are chosen using a circle with radius 2.5 times the Full Width at Half Maximum (FWHM) of a detected source, available from source detection algorithms and, for the corresponding background, an annulus with the same inner radius and an outer radius equal to $\sqrt{2}$ times of the inner one (it covers the same area in the sky as the source). It should be noted that the used radius for the source extracts the overwhelming majority of photon events corresponding to the source. If in the area used for the background, there is any source with maximum likelihood of existence greater than 0 $(ML > 0)$, the background is taken from the opposite direction of that source sector with 180° angle with inner radius equal to the $2.5 \times FWHM$ and outer radius $\sqrt{3}$ times of the inner one. Finally, as a background for a given source, a sector is used not including any source with $ML > 0$ and covering equal area in the sky as overwhelming majority of photon events coming from the source.

3. Example: An X-ray flare on G 131−026

In the direction of the PSPC pointing ID 700101 (P.I. Turner) with the UV Cet type star G 131−026 in the field of view, we identified a total of 117 X-ray sources above our detection threshold ($ML \geq 7.5$). All of the X-ray sources were cross-referenced with the SIMBAD and NED databases as well as with the UV Cet type flare stars and related objects catalogue (Gershberg et al. 1998). In addition, we inspected the HST Guide Star Catalogue and the Digitized Sky Survey images. In all cases, 40 arc sec circles around each X-ray source were chosen to identify possible optical counterparts to the X-ray sources (Neuhäuser et al. 1995).

The star G 131−026, also known as LTT10045, LP404-33, and CRSS J0008.9+2050, has a spectral type of M4.5e and is located at the distance of $11.3\,pc$ (Reid et al. 1995). We have identified it with the very strong X-ray source RX J0008.8+2050 ($ML > 10000$).

The procedure described above for variability testing has been applied to this source. Our method has divided the observation into six time segments in the broad energy band $(0.1\text{–}2.4\,keV)$, while the dataset of the corresponding background can be represented by one segment, i.e. no significant variations have been observed in the background at the same time. As seen from Fig. (1) during the observation this star has also shown another flare event possibly triggered by the first one with relatively smaller amplitude. It should be noted that the more powerful flare is observed during two successive observational intervals lasting 161 and 1420 seconds each. The time gap between them was equal to 110 sec. In order to estimate flare parameters we have converted the observed count rates to the fluxes and luminosities for each segment subtracting background and radiation of the star in the quiescent level (the leftmost segment in the second panel of Fig. 1).

The count rates of the flare in each energy band are expressed as

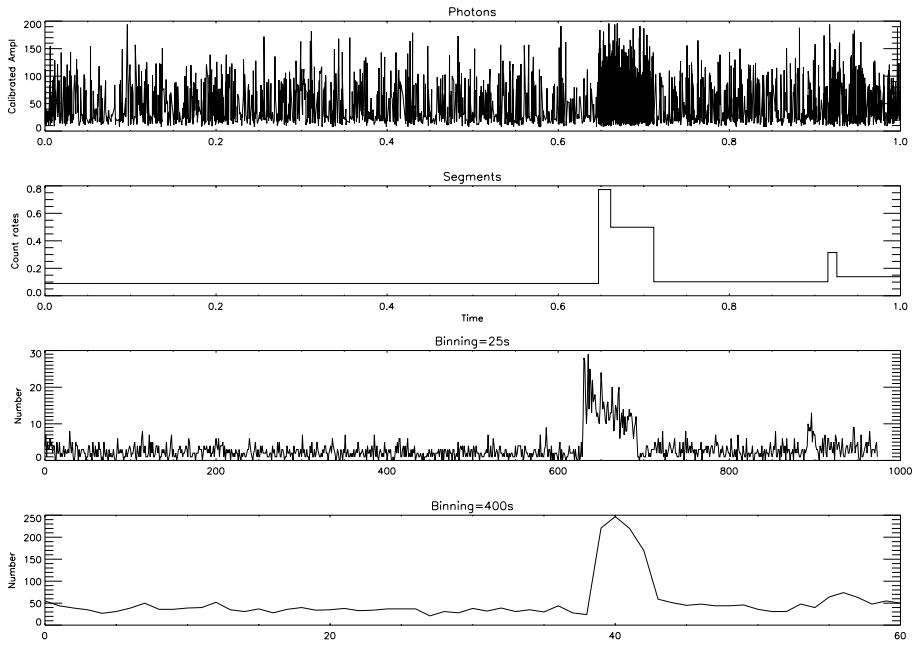

Fig. 1. *ROSAT* PSPC observations of the UV Cet type star G 131−026: In the upper panel, calibrated amplitudes (i.e. photon energy) of registered photons are plotted vs. arrival times. The ploting photon energy vs. arrival times and ignoring of observational gaps in the data set have been used for purposes of clarity. The second panel shows the Bayesian-blocks decomposition of these data. The next two panels show histograms of the same data binned into equal time intervals of 25 and 400 sec, respectively (except the last bin). In the last panel, the smaller flare near the end of the observation, is almost undetected, due to long time bins.

$$X_{Band}(t) = [X_{Flare}(t) + X_0]/X_0 - 1, \quad (5)$$

where X_0 is the quiescent star radiation count rates in the particular energy band, determined from data outside the flare event. This quantity can be used to determine the so called equivalent duration of the flare

$$X_{ED} = \sum X_{Band}(t_k)\Delta t_k, \quad (6)$$

where $X_{Band}(t_k)$ is given by Eq. (5) for each kth segment. $X_{Band}(t_k)$ expresses the energy of the flare in terms of the quiescent star's energy.

An estimate of the luminosity of the quiescent star, in each band, can be made by multiplying count rates by an energy conversion factor and taking into account the distance of the star. Assuming that the spectrum of the star in the quiescent level is consistent with a one-temperature Raymond-Smith spectrum (Raymond & Smith 1977), a thermal spectrum from an hot, optically thin plasma of solar abundance, we use 1 keV, i.e. $\sim 10^7\,K$, as temperature of the X-ray emitting plasma, which is typical for late-type stars (Neuhäuser et al. 1995). Because foreground absorption is negligible for our object, we use an energy conversion factor of $10^{-11}\ cts\ cm^2\ erg^{-1}$ (Zimmermann et al. 1998) for the PSPC observation.

Furthermore, flare detection and parameter estimation procedures have been applied to different energy bands. Namely, so called hardness ratios (X-ray colors) are defined as follows: If $X_{s,m,h}$ are the count rates in the bands soft (0.1 to $0.4\ keV$), medium (0.5 to $0.9\ keV$), and hard (0.9 to $2.0\ keV$), respectively, then

$$HR\,1 = \frac{X_h + X_m - X_s}{X_h + X_m + X_s} \quad \text{and} \quad HR\,2 = \frac{X_h - X_m}{X_h + X_m}$$

Ie., hardness ratios range from -1 to $+1$. If no counts are detected, e.g, in the soft band, then $HR\ 1 = \ 1$, but one can estimate a lower limit to $HR\ 1$ by using the upper limit to the soft band count rate X_s in the formula above.

These quantities have been computed for different observational phases, i.e. for the star in the quiescent level as well as for flare radiation alone. The results are presented in Table 1. The first column gives the Julian date at flare onset in different energy bands. Column 2 gives the energy bands at which the flare is detected (Broad: 0.1–2.4 keV, Soft: 0.1–0.4 keV, Medium: 0.5–0.9 keV and Hard: 0.9–2.0 keV). Column 3 is the normalized count rate at flare maximum as determined from Eq. (5). Eq. (6) was used to determine the equivalent duration in seconds, listed in Column 4. The count rates may be converted to the flare energies in different energy bands by multiplying by quiescent luminosity of the star given in Column 5 which is the logarithm of X-ray luminosity of the star at quiescent level and determined by the method outlined in this section. Columns 6 and 7 contain hardness ratios of the flare emission plus quiscent radiation, while the hardness ratios of the flare radiation and of the star in quiescent level alone are given in Columns 8,9 and 10, 11, respectively. These hardness ratios correspond to count rates observed in the segment including flare maximum and, for star only, the segment just before flare onset. The estimates of variations of hardness ratios are calculated by variations of count rates in each energy band. A more sophisticated treatment of best-fit parameter values and ranges is possible using full posterior probabilities distribution, but these simple intuitive estimates will be adequeate here. Obviously, the X-ray emission was harder during the flare, i.e. the plasma was hotter.

It is worthwhile to note that there is some time delay between the onset of the flare in different energy bands (see, Fig. 2). Namely, the flare radiation has started in the soft band, then in the medium band, and more later in the hard one. The ending of the flare was in the reverse order. Of course, this phenomenon might be different from one flare to another and a correct conclu-

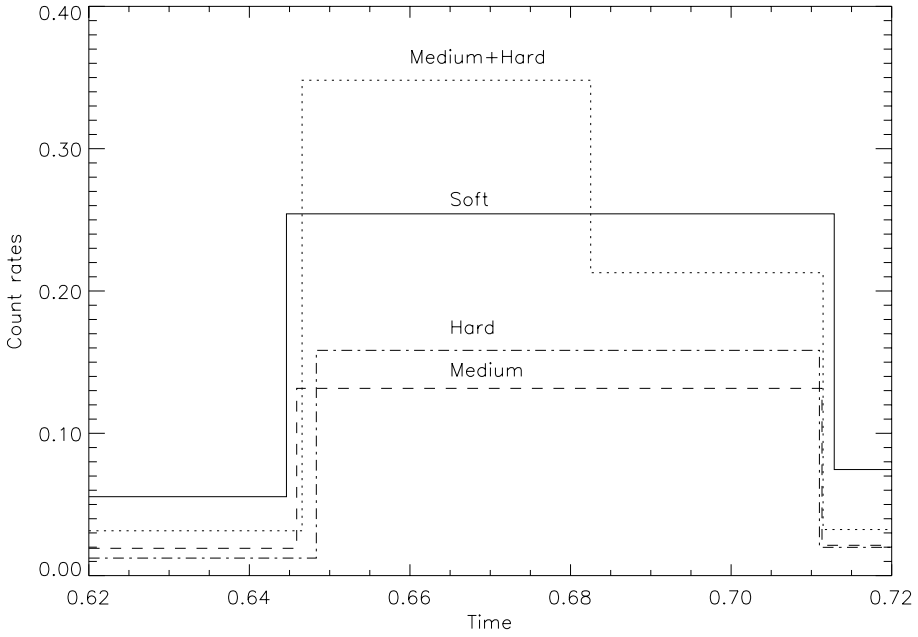

Fig. 2. Bayesian-blocks of the *ROSAT* PSPC observations of the UV Cet type star G 131−026 for observational intervals including the flare in different energy bands: Soft: 0.1–0.4 keV, Medium: 0.5–0.9 keV, Hard: 0.9–2.0 keV and Medium+Hard: 0.5–2.0 keV.

Table 1. Flare parameters for RX J0008.8+2050

JD (onset)	Band	X_{max}	X_{ED}	log L_X	Flare + Star		Flare Only		Star Only	
2448437.0000+			(sec.)	(quiet)	HR1	HR2	HR1	HR2	HR1	HR2
88564.5835	Broad	7.63	9076	28.08	−0.03	0.16	0.02	0.20	−0.32	−0.16
83994.6044	Soft	3.58	22417	27.85	±	±	±	±	±	±
84026.8215	Medium	5.84	36288	27.45	0.04	0.05	0.05	0.06	0.03	0.05
88570.3804	Hard	11.84	19650	27.25						

sion can be drawn only on the basis of statistics of significant number of flares which will be the subject of our subsequent paper (Hambaryan et al., in prep.).

4. Conclusions

Our analysis of *ROSAT* pointed observations shows that, for the detection of flares, it is straightforward to apply the method by Scargle (1998), namely partitioning the observational dataset into segments having statistically significant different count rates and using a formalism of Bayesian model comparison. The method is applicable to direct, unbinned data and in this sense it is not biased for determining the beginning and ending of flare events. It is very convenient to estimate the equivalent duration of the flare event, hardness ratios in different phases of flares as well as to detect any time delay in different energy bands.

Acknowledgements. VH wishes to acknowledge the warm hospitality of Professor J. Trümper and the Max-Planck-Institut für Extraterrestrische Physik during his stay as a Visiting Fellow. RN wishes to acknowledge financial support from the DFG Star Formation program. ROSAT is supported by the German government (BMBF/DLR) and the Max-Planck-Society.

References

Cash W., 1979, ApJ 228, 939
Cruddace R.G., Hasinger G.R., Schmitt J.H.M.M., 1988, In: Murtagh F., Heck A. (eds.) ESO Conference and Workshop Proc. 28, Astronomy from Large Databases. Garching, 177
Gershberg R.E., Katsova M.M., Lovkaya M.N., Terebizh A.V., Shakhovskaya N.I., 1998, Catalogue and bibliography of the UV Cet-type flare stars and related objects in the solar vicinity. A&A, in press
Haisch B. M., 1983, In: Byrne P.B., Rodono M. (eds.) Activity in Red Dwarf Stars. Astrophysics and Spoace Science Library 102, 255
Jaynes E.T., 1997, Probability Theory: The Logic of Science, available at http://bayes.wustl.edu
Jeffreys H., 1961, Theory of Probability. 3rd ed., Oxford University Press
Mirzoyan L.V., 1995, Optical flares: Observations and Interpretation. In: Greiner J., Duerbeck H.W., Gershberg R.E. (eds.) Proc. of IAU Colloquium No.151, Springer, p. 47
Montmerle T., 1997, In: Micela G., Pallavicini R., Sciortino S. (eds.) Mem. Soc. Astron. Ital. Vol. 68, No. 4, Soc. Astron. Ital., Firenze, p. 1017
Neuhäuser R., Sterzik M.F., Schmitt J.H.M.M., Wichmann R., Krautter J., 1995, A&A 297, 391
Pallavicini R., Stella R., Tagliaferri G., 1990, A&A 228, 443
Pfeffermann E., Briel U., Hippmann H., et al., 1986, Proc. SPIE 733, 519

Raftery A.E., 1994, Bayesian Model Selection in Social Research. University of Washington, preprint, available at http://www.stat.washington.edu/raftery/Research /Bayes/bayes_papers.html
Raymond J.C., Smith B.S., 1977, ApJS 35, 419
Reid I.N., Hawley S.L., Gizis J.E., 1995, AJ 110, 1838
Scargle J.D., Bapu G.J., 1998, Point Processes in Astronomy: Exciting Events in the Universe, preprint, available at http://ccf.arc.nasa.gov/∼scargle/hand_www.ps
Scargle J., 1998, ApJ 504, 405
Schmitt J.H.M.M., 1994, ApJS 90, 735
Trümper J., 1982, Adv. Space Res. 2 (no. 4), 241
Zimmermann H.U., Boese G., Becker W., et al., 1998, EXSAS Users Guide, MPE Report, *ROSAT* Scientific Data Center, Garching