DESY Summer student programme 2010

Statistical study of long term Gamma-Ray data

Supachai Awiphan

Department of Physics and Materials Science
Faculty of Science, Chiang Mai University, Thailand

Supervisors

Konstancja Satalecka

Elisa Bernardini

Multimessenger Young Investigator Group
DESY Zeuthen, Germany

Statistical study of long term Gamma-Ray data

Awiphan S.[1*], Satalecka K.[2], Bernardini E.[2]

[1] Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand
[2] DESY Zeuthen, Platanenallee 6, D-15738, Zeuthen, Germany

Abstract

In the following work, we carried out a statistical study of long term Very High Energy (VHE) Gamma-ray light curves of High frequency peaked BL Lac (HBL) sources (Mrk421, Mrk501, and 1ES1959 +650). We used the data collected from different experiments (Tluczykont et al. 2010) to generate long term light curves. During the analysis, we derived parameters and periodograms of the light curves. Using the results from our statistical study, we generated artificial continuous long term light curve and study the effects of different class of bias which are introduced by real observation constraints.

Key words : BL Lacertae objects, gamma rays

1 Introduction

Very High Energy (VHE, E>100GeV) Gamma-ray astronomy is one of the youngest branches of physics. It started around thirty years ago. In the first period of study in VHE Gamma-ray astronomy, the most effort was put into discovering new classes of sources or detecting the Gamma-ray signal from the known ones. However, the aim of study in recent year changed into analyzing the previously collected data in more detail.

Studying properties of sources in Gamma-rays has some problems, because the light curves[1] of sources in VHE Gamma-rays are not continuous like in other wavelengths. The observation gaps in data come from the limited duty cycle of the Imaging Atmospheric Cherenkov Telescopes (IACT) and restricted visibility periods of many sources. It's a limiting factor to any time variability analysis. In this study, we concentrate on studying the variability of Gamma-ray flux and generating Gamma-ray light curves of sources by using long term data from different IACT telescopes. The results will make us understand more clearly the variability of observed flux and the nature of the physical processes taking place in the source.

2 Blazars classification

An active galactic nucleus (AGN) is one type of extragalactic sources which we can observe in VHE Gamma-ray. This work on AGNs has concentrated on the most powerful class of AGNs called blazars. Blazars are often observable in all wavelength bands, from radio waves to Gamma-ray. In many of them, relativistic outflows (jets) are observed which are probably powered by mass accretion onto a supermassive black hole (10^6 - 10^{10} solar masses) in the galactic center. The characteristic property of blazars is that the jets are directed at a small angle with respect to the line of sight of the observer.

The radiation from blazars generally have a continuous Spectral Energy Distribution (SED) with two peaks, one from radio through optical to Ultraviolet range, is most likely due to synchrotron emission from relativistic electrons in the jet. The second emission component extends through X-ray and Gamma-ray energies and might be explained by the Synchrotron-Self Compton (SSC) model. This model assumes that the VHE Gamma-ray production of AGNs come from Compton scattering of lower energy radiation by the same relativistic electrons which are responsible for the synchrotron emission at lower frequencies. However, there are alternative models in which the high-energy emission is produced by hadronic processes in the jet.

High frequency peaked BL Lac (HBL) is a subclass of blazars which was first known to emit VHE Gamma-ray after the detection of Mrk421 above 300 GeV by the Whipple group. Moreover, the most frequently observed AGN sources that were early discovered belong to the HBL class.

*Corresponding author. E-mail: fluke_supachai@hotmail.com

[1] A light curve shows the flux measured from a given source as a function of time.

Due to early discovery, there is more VHE Gamma-ray data of collected from HBL sources than from any other subclass of blazars. In this work, we will focus on three HBL sources that were early discovered: Mrk421, Mrk501 and 1ES1959 +650.

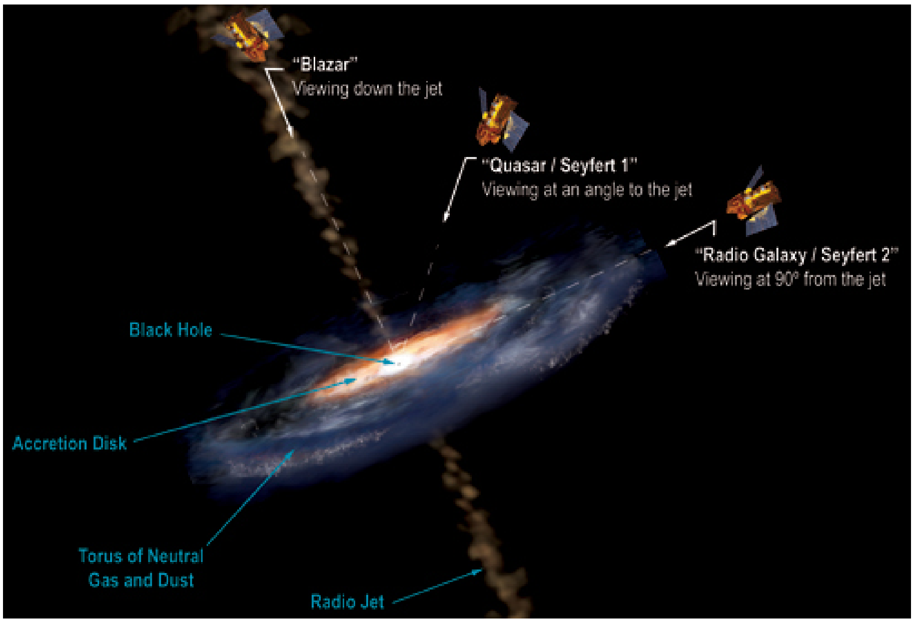

Figure 2.1: Features of an active galactic nucleus (AGN) and how to classify a type of AGN by the difference in viewing angle (Courtesy NASA).

3 Statistical analysis

3.1 Light curve

We use the Gamma-ray data collected by the multi-messenger group at Deutsches Elektronen-Synchrotron (DESY), Germany. In this work, we use the public light curve data from 1992 until today that were collected from Whipple, HEGRA, CAT, H.E.S.S., MAGIC, and VERITAS (Tluczykont et al. 2010) experiments. The data was converted to a common threshold of 1 TeV to generate long term light curves of three sources: Mrk421, Mrk501 and 1ES1959 +650. The upper limits were excluded from this study. The light curve shows the measured Gamma-ray flux integrated above 1 TeV as a function of time in Modified Julian Day (MJD)[2](Fig 3.1).

3.2 Flux state distribution

The distribution of the Gamma-ray flux values integrated above 1 TeV in one day bins for Mrk421 is shown in left panel of Fig 3.2. We described the flux state distribution by Gaussian and Log-normal distribution. Low state flux measurements (less than 0.5 Crab) are described by the Gaussian distribution. The mean of Gaussian distribution may represent the baseline flux. For high state flux measurement, the Log-normal distribution has been chosen. Log-normal variability is probably related to accretion disk activity. This type of variability has been discovered in X-ray emission of various compact systems: Seyfert galaxies, X-ray binaries and blazars (Giebel & Degrange 2009).

The probability density function of Log-normal distribution is defined by:

$$f(x:\mu_{Ln},\sigma_{Ln}) = \frac{1}{x\sigma_{Ln}\sqrt{2\pi}}\exp(-\frac{1}{2\sigma_{Ln}^2}(\ln(x)-\mu_{Ln})^2), \tag{3.1}$$

for $x > 0$, where μ_{Ln} is the mean and σ_{Ln} is the standard deviation of the Log-normal distribution.

[2]The Modified Julian Day (MJD) is an abbreviated version of the old Julian Day (JD), MJD=JD-2400000.5. The JD is measured from January 1, 4713 BC Greenwich noon.

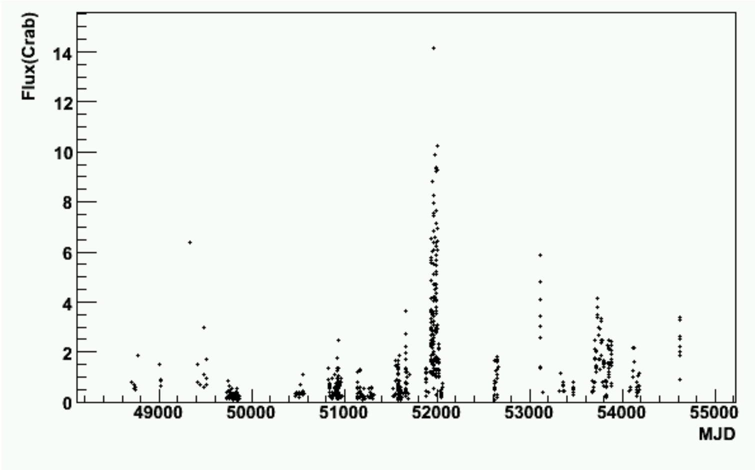

Figure 3.1: Long term light curve of Mrk421: in units of Crab (units of the flux observed from the Crab Nebula, the standard candle of Gamma-ray astronomy)

3.3 Probability of high states

The probability that the source in a flux state higher than a certain threshold, F_{th} is defined as:

$$P_{F>F_{th}} = 1 - \int_0^{F_{th}} f_{total}(x)dx, \tag{3.2}$$

where $f_{total}(x)$ is the function fitted to the flux state distribution (the sum of the Gaussian and Log-normal distribution). To find out the high state threshold, we use Gaussian distribution parameters from the fit. We define the source as being in high state when the flux level exceed the μ_{Gauss} by $N\sigma_{Gauss}$. Table 1 shows different threshold flux value for each source where N is 5, 10, and 20.

Table 1: The threshold of high state flux and probability that the source in the higher state than the threshold of 5, 10, and $20\sigma_{Gauss}$.

	Mrk421		**Mrk501**		**1ES1959 +650**	
	F_{th}(Crabs)	$P_{F>F_{th}}$(%)	F_{th}(Crabs)	$P_{F>F_{th}}$(%)	F_{th}(Crabs)	$P_{F>F_{th}}$(%)
$5\sigma_{Gauss}$	0.9	45	1.1	16	1.3	9
$10\sigma_{Gauss}$	1.5	31	1.9	10	2.7	2
$20\sigma_{Gauss}$	2.6	16	3.5	6	5.4	-

The probabilities of finding Mrk421 in high state at all flux threshold levels are higher than for other sources. From this result and flux state distribution, we can conclude that Mrk421 seems to have high states more often than other sources. Mrk501 show longer period of low state with occasional flare. For 1ES1959 +650, most of data that we have is in low state with only one prominent flare.

3.4 Delta flux distribution

We call "delta flux", the time derivative of the observed flux. To calculate the delta flux distribution, we use only two nearby flux measurements which have the start time of the observation not more than 2 days apart to avoid the long time gaps present in the data. The delta flux distribution of Mrk421 is shown Fig 3.2 (right panel). The distributions can be described by a Log-normal distribution. Moreover, we found that the changing in binning size of delta flux distribution do not affect the values of the fitted parameters. This is a confirmation that the original data should be properly described by a Log-normal distribution.

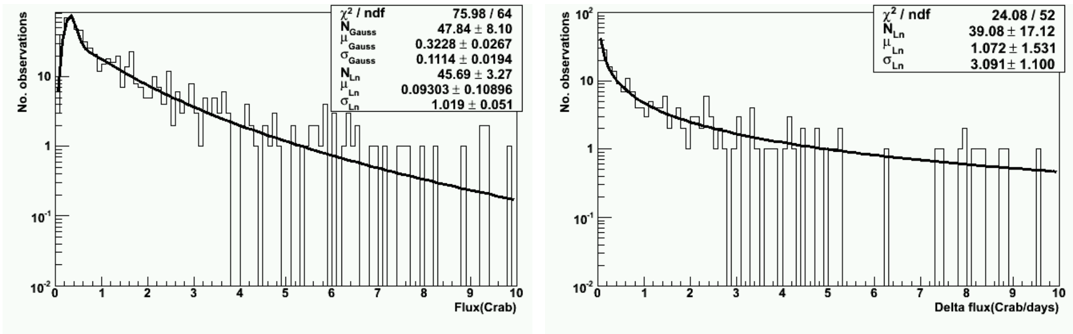

Figure 3.2: Flux state distributions of Mrk421 with fit function as a combination of the Gaussian and Log-normal distribution (left panel). Delta flux distributions of Mrk421 fitted with Log-normal distribution (right panel).

3.5 Relation between delta flux and flux

We plot delta flux as a function of flux to understand the fluctuations (Fig 3.3, left panel). A linear equation was used to fit the graph. The results show that the fluctuations are proportional to the flux level. The slopes that we got from fitting the three sources; Mrk421, Mrk501 and 1ES1959 +650 have values in the range of 0.25-0.50 Crab.

3.6 Relation between excess variance and average flux

The observed light curves will have finite uncertainties due to measurement errors. These uncertainties on the individual flux observations will contribute to an additional variance. This leads us to use the "excess variance" to estimate the intrinsic source variance (Vaughan et al. 2003). The excess variance is defined by:

$$\sigma_{xs} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2 - \sigma_i^2}, \tag{3.3}$$

where $x_i$is the i-th flux measurement, $\overline{x}$ is average flux and σ_i is statistical error of the i-th flux measurement. For this calculation, the data were divided into time-intervals of 50 days, each comprising at least 20 flux measurements. The excess variance estimates the variability corrected for the Poisson noise. In Fig 3.3 (right panel), we show the excess variance versus the mean flux of Mrk421. As can be seen, it confirms the result from the previous section, that the variability of the flux is directly proportional to the mean flux level.

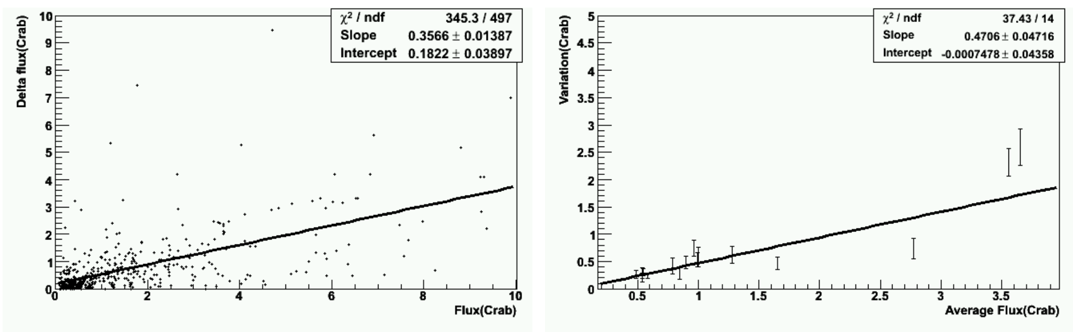

Figure 3.3: The delta flux versus flux of Mrk421 (left panel) and the excess variance versus the average flux of 50 days intervals with at least 20 flux measurements of Mrk421 (right panel). The black solid line represents a fit with linear function.

3.7 Periodogram

One of the most common tools for examining AGN variability is the Power Spectral Density (PSD) which represents the amount of variability power as a function of frequency. The high frequency PSDs of AGNs are usually well-represented by power-laws over a broad range of frequencies. The relation between power and frequency can be written as:

$$P(f) \propto f^{-\alpha}, \tag{3.4}$$

where $P(f)$ is the power at frequency f. The PSD can be estimated by calculating so-called the *periodogram*. For a light curve with period ΔT, the periodogram is the modulus-squared of the Discrete Fourier Transform (DFT) of the data. For a light curve comprising a series of fluxes x_i measured at discrete times t_i where $i = 1, 2, 3, ..., N$. The power at frequency v_p is:

$$I(v_{p_j}) = A|DFT(v_p)|^2 = A|\sum_{i=1}^{N} x_i e^{2\pi i f_j t_i}|^2 = A\{\sum_{i=1}^{N} x_i \cos(2\pi f_j t_i)\}^2 + A\{\sum_{i=1}^{N} x_i \sin(2\pi f_j t_i)\}^2, \tag{3.5}$$

where A is a normalization. We use $A = \frac{2\Delta T}{N}$ in order to give the periodogram in absolute units (Vaughan et al. 2003).

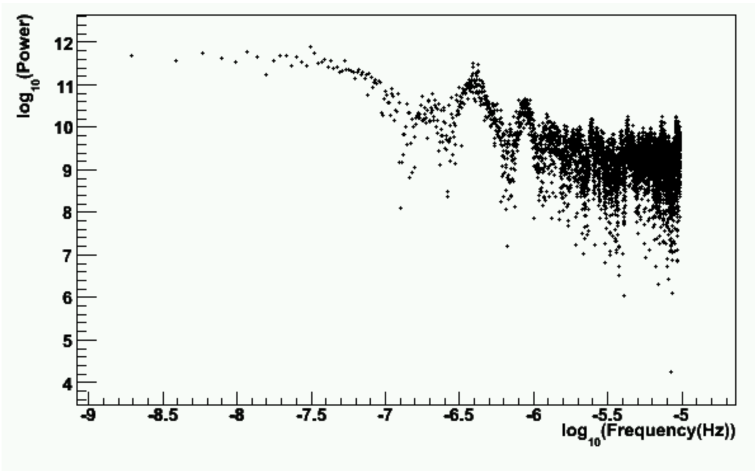

Figure 3.4: Periodograms of Mrk421.

The periodograms of Mrk421, Mrk501 and 1ES1959 +650 show some strange structures (peaks) in the frequency range 10^{-7} - 10^{-6} Hz. These kind of structures were not found in previous studies of periodograms of HBL sources in the X-ray bands. However, in a multi-wavelength study of Mrk501 by Rödig et al. 2002, they found that Mrk501 show a 72, 35 and 23 days period. At first, we assumed that the peaks in the periodograms may from same periodic nature of the Gamma-ray emission from this sources. We calculate the periods by fitting the periodograms with polynomial of 2^{nd} degree and using the first derivative to find out the position peaks. To find out the optimum range of fit, we compare the χ^2 per number of degree of freedom (NDF) value of each fit in the frequency range 10^{-7} - $10^{-6.4}$ for the 1^{st} period, $10^{-6.6}$ - $10^{-6.2}$ for the 2^{nd} period, and $10^{-6.3}$ - $10^{-5.9}$ for the 3^{rd} period. We use the range that yielded the smallest value of χ^2 per number of degree of freedom.

We compare the periods that we got for Mrk501 from our study with those from Rödig et al. (2002). We find that our periodicities have different lengths in comparison to previous studies, but they are still of the same order.

Our second hypothesis was that some of the structures came from the gaps the of observation time. We plotted the periodogram of Monte Carlo generated light curve which we sampled at the same observation times as in the original data. The sampling introduces peaks in periodogram, but the structures of the peaks are not exactly the same as in the real data.

Table 2: Periods from a fit of the periodogram with polynomial 2^{nd} degree function.

Period (Days)	**Mrk421**	**Mrk501**	**1ES1959 +650**
1^{st} period	58.9±0.8	64.6±0.7	58.8±0.4
2^{nd} period	29.0±0.1	27.46±0.05	25.58±0.02
3^{rd} period	13.13±0.01	16.93±0.04	13.13±0.03

3.8 The stationarity of light curves

The analysis of periodograms is meaningful only when the underlying processes are statistically stationary. We used two independent methods in order to test the stationarity of the VHE Gamma-ray light curves.

3.8.1 Comparing PSDs

First, we use the method proposed by Papadakis & Lawrence (1995) based on an original idea by Jenkins (1961). This method works by comparing the PSDs from different periodograms. If the PSDs show significant difference, the underlying process can be said to be strongly non-stationary.

In the first step, we divide the time series into two parts and calculate the periodogram of both parts, $I_{partI}(v_p)$ and $I_{partII}(v_p)$. We have to test this hypothesis at all frequencies. In order to do it, we calculated the following parameters:

$$S(v_p) = \frac{\log[I_{partI}(v_p)] - \log[I_{partII}(v_p)]}{\sqrt{\mathrm{var}\{\log[\mathrm{I_{partI}(v_p)}]\} + \mathrm{var}\{\log[\mathrm{I_{partI}(v_p)}]\}}}, \tag{3.6}$$

where $\mathrm{var}\{\log[I(v_p)]\} = \mathrm{var}\left\{\frac{\chi_2^2}{2}\right\} = \frac{\pi^2}{6(\ln 10)^2} \approx 0.310$ (Papadakis & Lawrence, 1993). We assume that the periodogram estimates are distributed as χ_2^2 variables. Then, the mean and variance of the random variable $S(v_p)$ have to be 0 and 1, respectively. Finally, we form the test statistic S:

$$S = \frac{1}{\sqrt{p_{max}}} \sum_{p=0}^{p_{max}} S(v_p), \tag{3.7}$$

here p_{max} is total number of frequencies over which we have estimated the periodogram. If p_{max} is large enough, S will be normally distributed with zero mean and unit variance.

The absolute values of S that we got were 33.4 and 27.5 for Mrk421 and Mrk501 respectively. For 1ES1959 +650, the number of measurement in second part is too low when compared with the number of data in the first part and we couldn't calculate its $S - Value$. S is expected to be normally distributed with zero mean and unit variance, but the result that we got from Mrk421 and Mrk501 $S - Value$ was very largely ($\approx$ 25-30 σ) away from zero. We conclude that the light curves of Mrk421, Mrk501 and 1ES1959 +650 are strongly non-stationary.

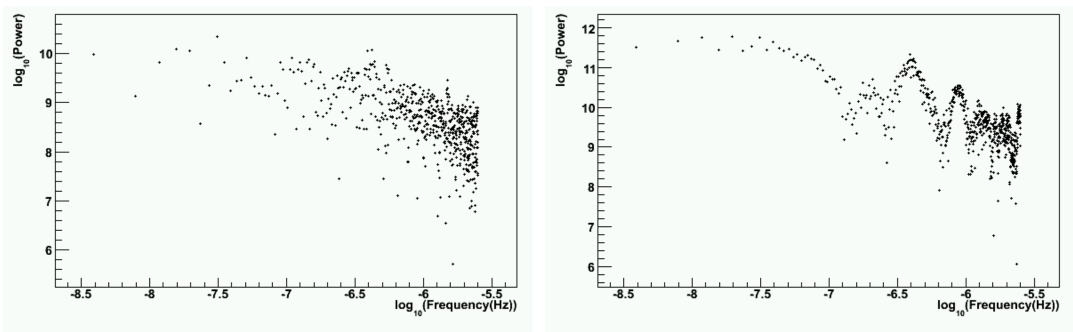

Figure 3.5: Periodograms of first (left panel) and second (right panel) parts of the light curves that used to find $S - Value$ of the light curves of Mrk421.

In Fig 3.5, the periodograms of the first and second parts of the light curve are shown. Surprisingly, only one of the periodograms from each pair shows strange structures (peaks) - the part which contains a large flare. From this result, we assume that the large flare generated strange structures on periodograms of all sources and made their light curve look non-stationary.

3.8.2 Comparing variances

In the second method (Vaughan et al. 2003), the variance of the observed flux is used to test the stationarity of the light curve. In Fig 3.6, we show the flux variances as a function of time to see the fluctuations in the light curve. Those variances were calculated for 50 days intervals with at least 10 flux measurements each. The confidence intervals were calculated using the standard deviation of the variances.

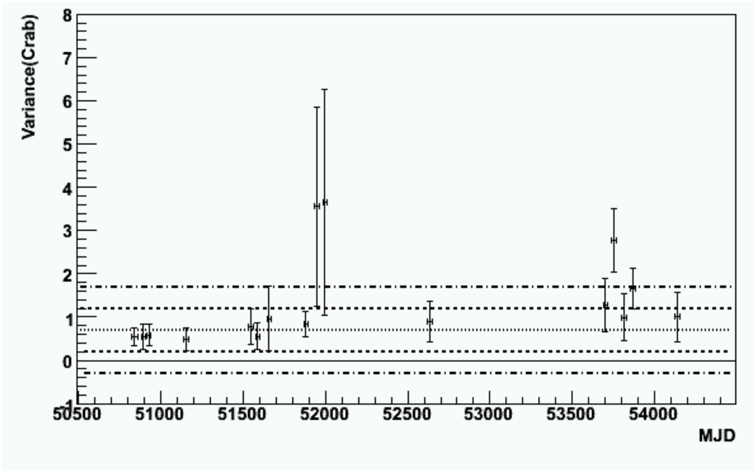

Figure 3.6: Variance of the flux avaeraged over 50 day intervals with at least 10 measurements as a function of time for Mrk421. The lines show with average variance (dotted line), 1σ (dashed line) and 2 σ (dash-dotted line) confident interval.

We can see that all sources have at least one variance value lying outside of the 2σ confidence interval. However, if we look on the error bar of the variances, we find that almost all variance values, including those from the flaring period are compatible with 2σ confidence level. Furthermore, we have a small number of points in these plots. Then, we cannot conclude that the light curves of Mrk421, Mrk501 and 1ES1959 +650 are stationary or not. In the future work, the confidence interval of variance distribution should be calculated using a Monte Carlo simulation like in Vaughan et al. (2003) for stronger conclusion

4 Generating a "realistic" VHE Gamma-ray light curve

4.1 Algorithms to simulate a light curve from the flux state distribution

As discussed in Section 3.2, we can fit the flux state distribution with Gaussian and Log-normal distribution. Then, we use the parameters of Gaussian and Log-normal distribution obtained from the fit to generate VHE Gamma-ray light curves of HBL sources.

4.1.1 Random method

The first method that we explored to generate artificial light curves is a random method. In this method, we just randomly generate flux from Gaussian and Log-normal distributions with parameters obtained from the fit to the distributions of collected data. The light curves that we got from this method (Fig 4.1, left panel) have the same flux distribution as original data, but the evolutions of flux level have a different structure and the artificial light curve looks too "spiky". The absolute values of the flux difference on two consecutive days are higher than the original ones.

4.1.2 Rearrange method

To solve the problem of high value of flux fluctuations of the results from simulation obtain in the random method, we introduced a time evolution constraint. In this algorithm, we rearrange the generated fluxes in ascending order. The flux value generated in the n-th step is then taken from the values with indexes in the range of $n \pm 10\%n$. The result in Fig 4.1 (right panel) shows the light curve generated according to this method. We can generate light curves with combined low state and flares. However, we don't know how to realistically model the expected time evolution.

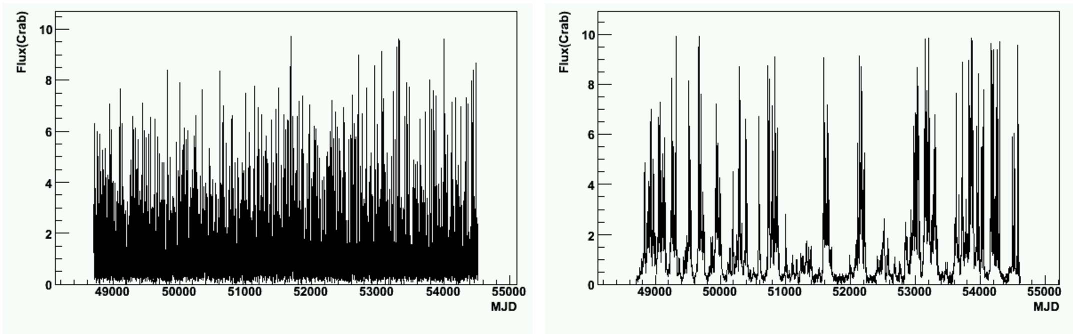

Figure 4.1: Light curve simulated using the random method (left panel) and rearrange method (right panel) and the flux state distribution of Mrk421.

4.2 Algorithm to simulate a light curve from the delta flux distribution

In section 3.4, the delta flux distribution was introduced. The parameters of from a Log-normal fit to the data are used to generate flux values to constraint the time evolution of generated light curves. We generated random delta flux from the distribution and add it to the flux in pervious step of simulation. An example light curve generated with this algorithm is shown in Fig 4.3. The result shows flares of long duration with very high flux (more than 100 Crab) which don't happen in the original data.

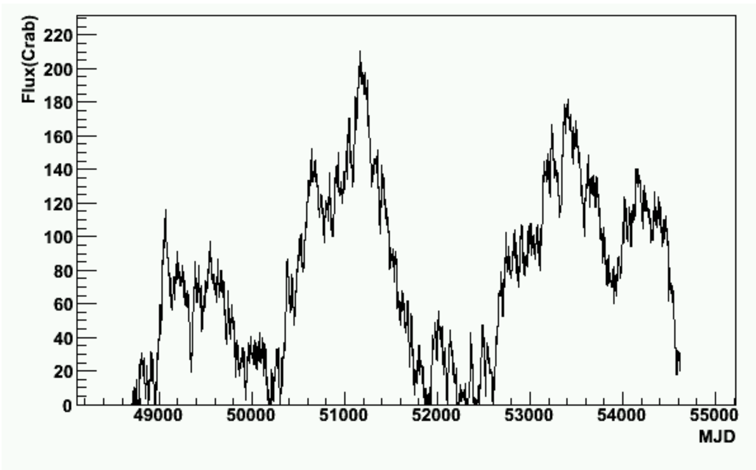

Figure 4.2: Light curve simulated using the delta flux distribution of Mrk421.

4.3 Algorithm to simulate a light curve from periodogram

To generate light curve from periodogram, we use the algorithm proposed by Timmer & König (1995). This algorithm is based on a main result of the theory of spectral estimation, i.e. that $DFT(v_p)$ is a complex is Gaussian random variable:

$$DFT(v_p) = \aleph(0, \frac{1}{2}S(v_p)) + i\aleph(0, \frac{1}{2}S(v_p)), \tag{4.1}$$

where $S(v_p)$ is the data frequency spectrum, which variance does not depend on the number of data points. These random variables are uncorrelated for different frequencies.

In the first step of this algorithm, we have to choose a power law $S(v_p) \sim (1/2\pi v_p)^{\alpha}$ from which we want to generate the light curve. For each Fourier frequency v_p, we have to generate two Gaussian distributed random numbers and multiply them by $\sqrt{\frac{1}{2}S(v_p)} \sim (1/2\pi v_p)^{\alpha/2}$. Then we use the result as the real and imaginary part of the Fourier transform in Equation 4.1. After that, we have to obtain the time series by using Inverse Discrete Fourier Transform (IDFT) of $f(v_p)$ to convert the values from the frequency domain to the time domain.

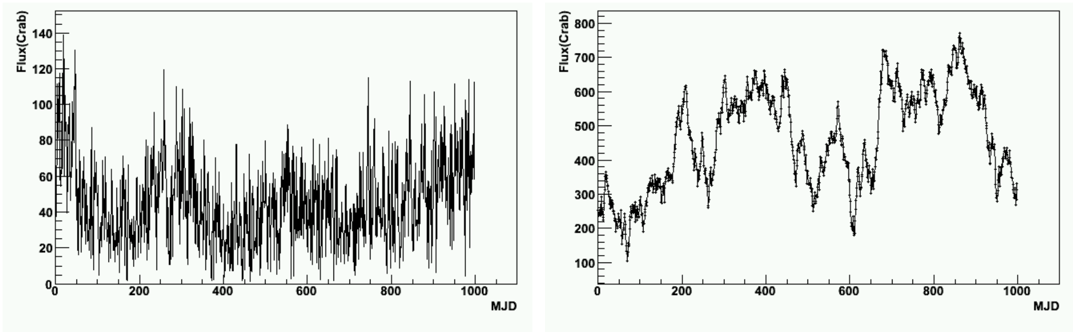

Figure 4.3: Simulated light curves from periodogram of α =1 (left panel) and 2 (right panel).

Examples of simulated light curves with this algorithm are shown in Fig 4.3. For $\alpha = 1$, we can simulate flicker noise light curve. In this case, the first frequency bin contributes the largest part of the variance. This algorithm ensures that the first frequency bin does not dominate the time series in a deterministic manner, but according to its natural fluctuations. To simulate random walk light curve, we have to choose $\alpha = 2$. Compared to the flicker noise light curve, the random walk light curve is dominated by longer timescales. When we calculate the periodogram of generated light curves, we found that the slope of the periodograms have increased in comparison with the input α. This problem may came from using estimate value of $\sqrt{\frac{1}{2}S(v_p)} \sim (1/2\pi v_p)^{\alpha/2}$ to generate light curve.

4.4 Algorithm to simulate a light curve from periodic function

Gaskell (2004) suggested that there is no fundamental difference between high state and low state and division between high state and low state is not significant. They used sine wave to describe the mean flux level and fluctuation proportional to that mean flux level to generate light curve. We apply this suggestion to implement a new algorithm to simulate light curves. The sine wave function with a period of 100 days is used to generate mean flux level and then we apply the flux fluctuations on top of it. To calculate the fluctuations Δx, we use the following relation:

$$\Delta x = \pm(S_d x + I_d) \pm Random\{0, (S_v x + I_v)\}, \tag{4.2}$$

where S_d, I_d, S_v and I_v are slope and intercept of linear functions describing the delta flux as a function of flux (Fig 3.3, right panel) and excess variance as a function of average flux (Section 3.6), respectively. The S_v and I_v are used to set the range of random flux fluctuations above the value which we get from S_d and I_d. The plus-minus signs are also set by random.

4.5 Algorithm with separate high state and low state

4.5.1 Generating a big flare and a low state from the periodogram

In section 3.8.1, the periodograms of two parts of real light curve show that the strange structures in periodogram may come from a part with flares. Moreover, the result indicates that the light curves may not be stationary. Both conclusions indicate that we may not be able to generate the flares and low state in the same way. Then, we separate the data into two parts: a big flare and a low state. To define a big flare, we use the 100 days interval that includes the highest flux of the light curve and the other periods we define as low state.

The periodograms from original data for both parts were plotted and fitted with a power law (Fig 4.4). We used different slops from the power-law function for a big flare and low state used to generated light curve with our algorithm as presented in section 4.3.

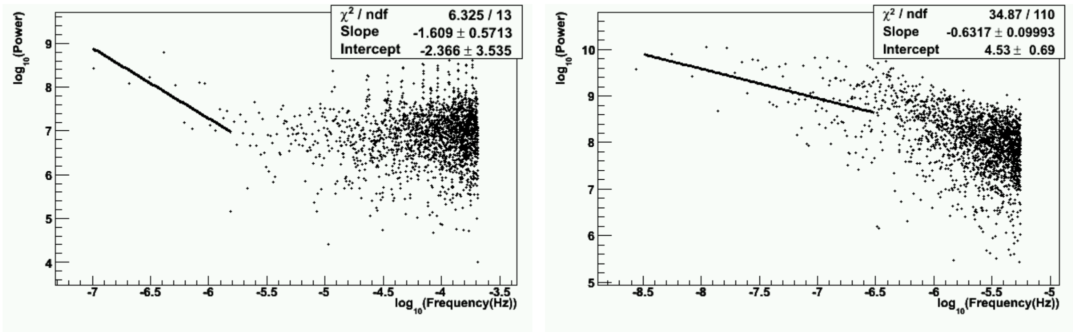

Figure 4.4: Periodograms of a big flare (left panel) and low state (right panel) of Mrk421.

Fig 4.5 shows the simulated light curve and corresponding periodogram with parameters for Mrk421. There is a problem on the transition between low state and a big flare that makes our light curve look like a step function. This problem comes from our algorithm that doesn't have a proper transition function between low state and a big flare.

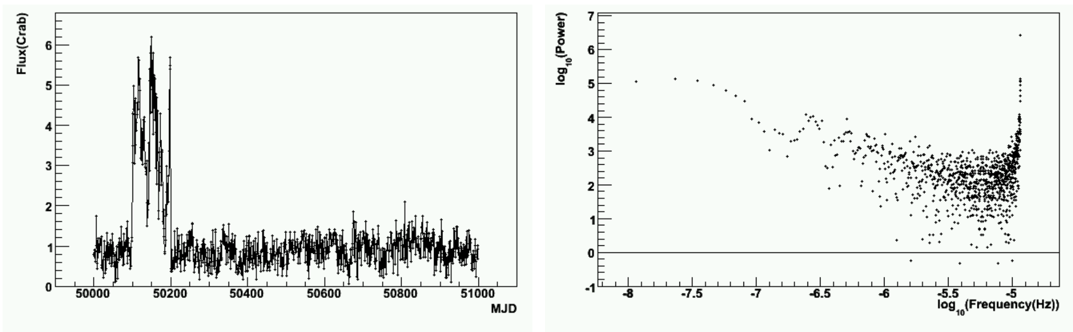

Figure 4.5: Light curve with a big flare and low state simulated from periodogram (left panel) and corresponding periodogram (right panel). Parameters derived from Mrk421 were used for simulation.

4.5.2 Generating a big flare from a periodic function and a low state from periodogram

From section 4.5.1, we have a problem with transition between low state and big flare. In section 3.7, we found that the light curve of HBL source have periods which we can find by fitting the peaks in periodograms with a polynomial of 2^{nd} degree. To make the transition between a big flare and low state, we apply the algorithm that generates a light curve with sine wave to generate a big flare. The sine wave function that we use is:

$$x_i = (\overline{x_F} - \overline{x_L}) \sin(\frac{2\pi t_i}{T} - \frac{\pi}{2}) + \overline{x_F}, \tag{4.3}$$

where $\overline{x_F}$ and $\overline{x_L}$ are the average flux of a big flare and low state, respectively, and T is the period of a sine wave. Light curves from this simulation have smooth transition between a big flare and low state. Same as in section 4.4, we add random flux fluctuation on top of the sine wave function. We apply this simulation to simulate the light curve with periods of sine wave as obtained from real data and then we sample the generated light curve using the observation pattern of the real data (Table 2). The shape of the resulting periodograms is very similar to the real ones.

However, in Table 2, we have three periods per sources which we obtain from fitting the periodogram. We want to find which period of sine wave function produces a simulated light curve

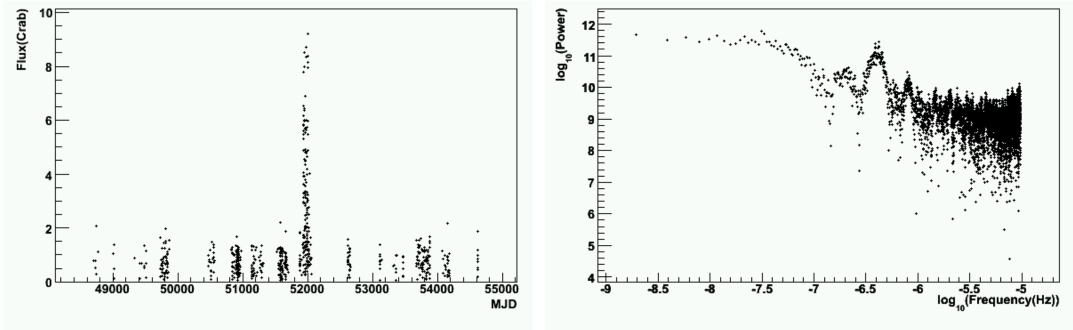

Figure 4.6: Simulated light curves (left panel) and corresponding periodograms (right panel) of Mrk421. The big flare is generated from periodic function and low state flux is generated from periodogram.

most similar to the original one. To check this, we compared the PSDs (Section 3.8.1). We simulated 1,000 light curves from this algorithm and found the average *S-Value* between them and the periodogram of original data.

Table 3: Average *S-Value* for different high state period values (see Table 2) calculated from 1,000 generated light curves

S-Value	Mrk421	Mrk501	1ES1959 +650
1^{st} period	6.0	1.5	14
2^{nd} period	0.8	1.2	11
3^{rd} period	1.7	1.2	20

The average *S-Values* that we obtain in Table 3 show that using second period to generate big flare gives the best average *S-Values* for all sources. For Mrk421, we obtained a very similar light curve compared to the original one when we used the second period (the *S-Values* is less than 1σ). The simulated light curves with the second period of Mrk501 are quite the same as the original light curve, but have higher value of average *S-Values* than Mrk421 (1σ-2σ). In the case of 1ES1959 +650, the values that we got are very high (more than 10σ). The very high value of average *S-Values* may come from that we have few measurements and most of the data of 1ES1959 +650 have the same flux value and the flux of a big flare doesn't differ from the low state so much.

4.5.3 Generating high state from periodic function and low state from periodogram

In section 4.5.1 and 4.5.2, we simulated the high state with only one big flare, but in real data, we can observe also smaller sub-flares in the big flare and single small flares. In order to reproduce those structures we have to answer two following questions. First, how many sub-flares a big flare has? And second, is there a relation between the duration of a flare and its maximum flux?

We set a threshold on the flux above which we define the source as being in a high state and then measure how long are these periods of high state and how many flares occur. In this work, we define as high state flux which exceeds the μ_{Gauss} by $20\sigma_{Gauss}$ (from the fit described in Section 3.3). We take the second period from Table 2 to define the standard duration, called "cycle", of a flare and calculate how many cycles are observed during the high state period in real data.

We describe the histogram of number of cycles by an exponential distribution. For each flare, we calculated average flux and maximum flux to plot them as a function of the number of cycles and fit with linear relation. The result in Fig 4.7 shows that average flux and maximum flux in each flare are proportional to the number of cycles.

As in the previous section, we generated flares with sine wave function. However, in this simulation, the sine wave function depends on more parameters:

$$x_i = (N_C S_{max} + I_{max} - N_C S_{avg} - I_{avg}) \sin(\frac{2\pi t_i}{T} - \frac{\pi}{2}) + (N_C S_{avg} + I_{avg}), \quad (4.4)$$

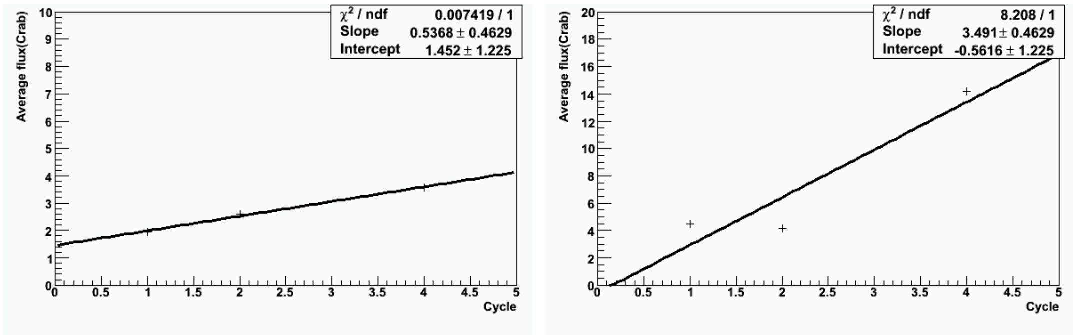

Figure 4.7: Average flux (left panel) and maximum flux (right panel) of high state as a function of the number of cycles, in each flare of Mrk421.

where S_{max}, I_{max}, S_{avg} and I_{avg} are the slope and intercept of the linear fit of Fig 4.7, respectively, and N_C is integer number of cycles generated randomly from the exponential distribution.

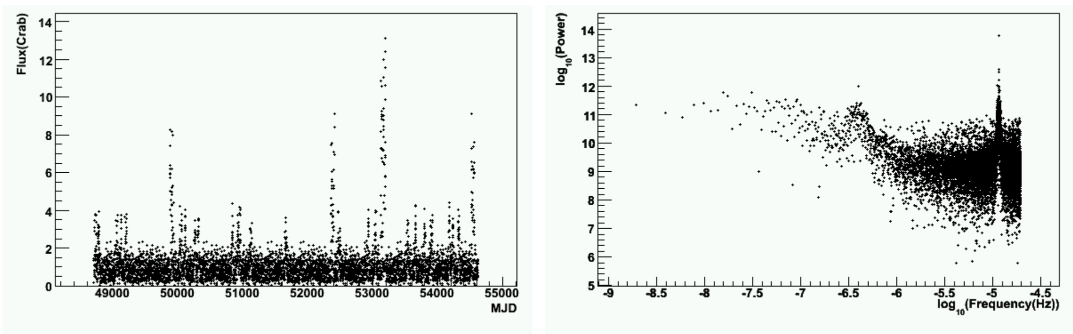

Figure 4.8: Simulated light curve. High state are generated from periodic function and low state are generated from periodogram (Top) and periodogram (bottom). Simulation based on parameters derived from that data from Mrk421.

In Fig 4.8, we show a simulated light curve from this algorithm, and corresponding flux distribution and periodogram. The flux state distribution well reproduces the original data of Mrk421. For the periodogram, it does not show strange structure like the original one. A peak at $10^{-6.4}$ Hz comes from the period that we use to generated high state (29.0 days) and a peak with frequency $10^{-4.9}$ Hz relate to a day since the light curve were generated with one flux point for each day. In this simulation we did not apply the observation time pattern of real data. Therefore, we conclude that this bias may have a stronger effect on the periodogram, compared to the flares.

5 Conclusions

A statistical study of long term VHE Gamma-ray light curves of three HBL sources (Mrk421, Mrk501, and 1ES1959+650) was performed. The data was previously collected from different experiments since 1992. We investigated the distributions of the integrated flux values and the distribution of the time derivative of the observed flux. We found out that the first one is well describe by a sum of a Gaussian and Log-normal distributions and the second one by a Log-normal distribution.

To investigate the frequency domain of the light curve, periodograms are produced by using Discrete Fourier Transform. The periodograms show strange peaks at frequency range 10^{-7} - 10^{-6} Hz. By fitting those peaks with polynomial of 2^{nd} degree, three periods are obtained for each source.

We tested the stationarity of the light curve with two methods: comparing PSDs and comparing variances. In the first method, the results show that all light curves are non-stationary. In the second test, due to small amount of data the results are non-conclusive.

In this study many different algorithms to generate light curves were developed and tested. They are based on the results from our statistical studies. First, we proposed an algorithm based on the flux state distribution and delta flux distribution, but they fail to properly describe flares. In the second approach, the periodogram and periodic functions were studied to generated light curves. The light curves simulated with this algorithm can very well reproduce the behavior of the real light curve.

The algorithm that we use to generate light curve by using periodogram still have some problems that come from estimation of the power. In the future work, we will study this more detail. Also, the sine function does not exactly describe the structure of the flares. Exponential functions may be more appropriate to describe the structure of flare in future work. We intend to continue this work to understate the behavior of the VHE Gamma-ray sources more clearly.

6 Acknowledgements

It is my great honor to be one of the Thailand representatives participating in DESY summer student programme 2010 under the patronage of HRH Princess Maha Chakri Sirindhorn.

I would like to express my deepest gratitude and indebtedness to my supervisor, Konstancja Satalecka, for her support and guidance throughout the DESY summer student programme 2010. Her continued support led me to the right way. I would also like to extend my appreciation to my group leader, Elisa Bernardini, for her very valuable suggestions.

My sincere appreciation is extend to Astropartical Physics group people for their kindness and sympathy and special thanks to Karl Jensen and people who are organized this summer student program in providing me the tremendous opportunity of participating in the program.

I also wish to thank all students of DESY summer student program who are very friendly for their friendship and other DESY Zeuthen people especially Sakhon Rimjeam and Jatuporn Saisut ,Thai post-doc and PhD student at DESY Zeuthen respectively, for their help during summer student program.

Furthermore, Deutsche Elektronen-Synchroton (DESY, Germany), The National Science and Technology Development Agency (NSTDA, Thailand) and Synchrotron Light Research Institute (SLRI, Thailand) are thanked for their support.

7 References

[1] Bevington P.R., & Robinson D.K., 1992, *Data Reduction and Error analysis for the Physical Sciences*, McGraw-Hill (New York)

[2] Gaskell C.M., 2004, AJ, 612, 24

[3] Giebels B., & Degrange B., 2009, A&A, 503, 797

[4] Papadakis I.E., & Lawrence A., 1993, MNRAS, 261, 612

[5] Papadakis I.E., & Lawrence A., 1995, MNRAS, 272, 161

[6] Rödig C., Burkart T., Elbracht O., & Spanier F., 2009, A&A , 501, 952

[7] Timmer J., & König M., 1995, A&A, 300, 700

[8] Tluczykont M., Bernardini E., Satalecka K., Shayduk M., & Kalekin O., 2010, A&A (Accepted)

[9] van der Kils M., 1989, ARA&A, 27, 517

[10] Vaughan S. Edelson R., Warwick R.S., & Uttley P., 2003, MNRAS, 345, 1271