Long-Term Photometric Variability Analysis Methods for Quasars: Postprint

Authors: Yu Lian, Zhang Xiong, Zhang Haojing, Li Futing, Xu Xiaolin, Ren Guowei, Wu Yuecheng

Date: 2018-09-13T00:00:00+00:00

Abstract

This paper presents four methods for analyzing long-period optical variability of quasars. Using a simulated periodic signal y=sin to test these four analytical methods, the results indicate: 1. when the number of data points sampled from celestial optical variability is relatively small, below a certain threshold, the analysis results from the Jurkevich method, the Date-Compensated Discrete Fourier Transform method (DCDFT), the Discrete Correlation Function method (DCF), and the Power Spectral Density method (PSD) differ; after obtaining the shortest continuous data points, the Jurkevich method may provide the most accurate and reliable results among the four methods, and its computational approach is concise and practical. 2. The optimal parameters for the Jurkevich analysis method were obtained, with m=9 yielding the best analysis results. 3. Using the Jurkevich method with m=9, the optical variability periods of the quasar 3C 279 and the quasar 3C 454.3 were analyzed, yielding a possible optical variability period for 3C 279 of , and for 3C 454.3 of 457d.

Full Text

Study on Long-Period Variability Analysis Methods of Quasars

Yu Lian, Zhang Xiong†, Zhang Haojing, Li Futing, Xu Xiaoling, Ren Guowei, Wu Yuecheng
(College of Physics and Electronic Information, Yunnan Normal University, Kunming 650500, China)

Abstract

This paper presents four methods for analyzing long-period variability of quasars. Using a simulated periodic signal $y = \sin\theta$ to test these four analytical methods, the results show: (1) When the number of sampled data points for a celestial light curve is relatively small to a certain value, the analysis results from the Jurkevich method, the Date-Compensated Discrete Fourier Transform (DCDFT) method, the Discrete Correlation Function (DCF) method, and the Power Spectral Density (PSD) method differ. After obtaining the shortest continuous data points, the Jurkevich method’ s analysis results are likely the most accurate and reliable among the four methods, and the calculation method is simple and practical. (2) The optimal parameters for the Jurkevich analysis method were obtained, with the best analysis results when $m = 9$. (3) Using the Jurkevich method with $m = 9$ to analyze the light variation periods of the quasar 3C 279 and the celestial object 3C 454.3, the possible light variation period of 3C 279 is and that of 3C 454.3 is 457d.

Keywords: Quasars; Jurkevich method; DCDFT method; DCF method; PSD method; Long-period variability

Long-period variability is one of the important methods for studying the properties of blazars, and its determination relates to the estimation of multiple physical quantities associated with the structure and radiation of blazars, such as the radius of the radiation region, the Doppler factor of the jet, and the black hole mass. Long-period variability provides certain parameters for the establishment of theoretical models. Light variations can be divided into long-timescale, medium-timescale, and short-timescale variability [1]. For long-timescale and medium-timescale variability, due to various factors such as observational instruments, lunar phases, and weather, it is difficult to obtain relatively complete observational data sequences of light curves [2]. When studying the periodicity of optical violently variable quasars, by investigating long-period variability analysis methods, one can obtain optimal parameters for periodic variability analysis and obtain the best period estimates from limited actual observational data sequences.

Common period analysis methods include power spectrum, Jurkevich, wavelet, structure function, etc. These methods can reliably analyze uniformly sampled data above the Nyquist sampling limit [3-5]. However, in actual data analysis, particularly for long-timescale periodic analysis of blazars obtained from astronomical observations, the use of these methods is subject to many constraints. For example, Fourier analysis [6] requires continuous, equally spaced sampling, and the handling of missing data points introduces some unrealistic information. Therefore, applying these methods to periodic analysis of celestial light curves increases the error in period determination. The Jurkevich method, proposed by Jurkevich et al., is a statistical method based on expected mean square error that is very helpful for handling unequally spaced observational data [7]. This paper will discuss this issue using four methods and propose corresponding improvements to make the analysis and calculation of periodicity in quasar light curves with unequally spaced time series simpler and more accurate.

1 Calculation Methods for Long-Period Variability of Quasars

1.1 Jurkevich Method

The Jurkevich method is a statistical approach proposed to address non-uniform measurement problems in astronomical observations [7]. Literature [8] has used this method to analyze the variability characteristics of the BL Lac object PKS 0735+178. Assuming our observational sample consists of N data points, where X_i represents individual measurements, $\bar{X}$ is the mean of all measurements, V^2 is the sample variance, and S is the sample standard deviation, we have: . (3) , (1)

If the sample is divided into m groups, the statistical parameters for group l should be:

The total variance corresponding to m groups is: $2 <$. The above formula actually indicates that when n samples with the same independent variable are calculated, they become one sample. If the group means are equal, the final expected variance will equal the average variance of the groups. If the group means differ, the final expected variance will exceed the average variance of the groups [7].

In period analysis, we fold the data according to different trial periods P, then arrange the folded data in order of increasing phase and divide it into m groups. Given an observation time t and trial period P, the corresponding phase is defined as [7]: . (1)

where $\phi(t)$ satisfies mV. If the obtained experimental period equals the actual period, the time corresponding to the minimum value of $2mV$ in the 2 relationship curve can yield the sample period. t_0 represents the time origin. Calculating the variance of each group' s data, the value of $2mV$ will reach its minimum, determined by the trial period lV and the relationship between total variance nP and 2.

1.2 Date-Compensated Discrete Fourier Transform Method

The Date-Compensated Discrete Fourier Transform method is one of the most commonly used methods for calculating light variation periods, first proposed by Ferraz-Mello in 1981 [9]. In previous research, literature [10] used this method to analyze the infrared light variation period of PKS 1510-089. By performing Gram-Schmidt orthogonalization on $\cos(\omega t)$ and $\sin(\omega t)$, three orthogonal vectors are obtained, and the data is projected onto these three orthogonal vectors to obtain the spectrum. The specific process is as follows:

After orthogonalization of H_0, H_1, H_2:

The parentheses denote the inner product of two vectors. Then h_0, h_1, h_2 are determined according to the following relationships:

In the case of non-uniform sampling, the weighted Date-Compensated Discrete Fourier Transform (DCDFT):

In many quasar observations, the precision of observational data f_i varies. Considering this issue, weights w_i are introduced. After redefining the inner product with weights:

After introducing weights into the inner product, the regression coefficients are obtained:

The intensity at frequency ω is given by:

From linear regression theory, it is known that . Using this property, a normalization factor is introduced: the statistic . For all frequencies ω, we call this quantity the spectral correlation coefficient.

1.3 Discrete Correlation Function Method

The Discrete Correlation Function method (DCF) was developed by Edelson and Krolik (1988) to analyze the correlation between two sets of discrete data, and literature [11] used this method to analyze the variability of PKS 2155-304. This method can indicate the correlation and time delay between two variable time series and can be applied to periodic analysis [11]. The specific steps are as follows:

First, calculate the discrete correlation function values for two data sets. If arrays a and b are the two data sets, $\bar{a}$ and $\bar{b}$ are their respective means, and σ_a and σ_b are their respective standard deviations.

Second, calculate the DCF(τ) values. The two data sets are related through time delay $\tau \pm \Delta\tau/2$. If the time delay is τ and falls within the interval Δt_{ij}, then the DCF(τ) value is:

The error of the discrete correlation function is again . For the obtained discrete correlation plot, if the peak is to the right of zero, it indicates that array a changes later than array b. Conversely, if the peak is to the left of zero, it indicates that array a changes earlier than array b.

1.4 Power Spectral Density Method

The definition of power spectral density [12] is that if $u(t)$ is a function that can be Fourier transformed, then . Because $u(t)$ is a real function, $u(\nu)$ is generally a complex function, and they satisfy Parseval’ s formula:

If $u(t)$ represents the spectrum, the left side of the equation represents the total energy of $u(t)$ over $(-\infty, \infty)$. The integrand on the right side is called the energy spectral density, which is a non-negative real number representing the energy per unit frequency. In practice, most $u(t)$ cannot be Fourier transformed in the mathematical sense. If a truncated function $u_T(t)$ is used to 截取 $u(t)$, then for the truncated function $u_T(t)$ with finite duration, a Fourier transform can be performed:

Similarly, it also satisfies Parseval' s formula:

Relating to the truncated function, dividing both sides of the above equation by $2T$ and letting . Corresponding to the definition of energy spectral density, the integrand on the right side is called power spectral density, denoted as:

From the expression, it can be seen that power spectral density is a non-negative real number. From the entire derivation of power spectral density, it is evident that power spectral density is a quantity in the frequency domain, directly reflecting the values corresponding to different frequencies in the frequency domain.

2 Simulation Tests and Parameter Optimization

2.1 Simulation Test Results for Astronomical Periodic Signals

To test the reliability of the above four research methods, we used a simulated periodic signal as astronomical observational data to analyze these four methods. Here we use a sine function with period $\pi 2$ as the test signal, with units in radians. In the experiment, we selected rad0 as the starting point, with a step size of rad1.0. A total of 15 groups with different numbers of data points were taken, and all research methods considered factors such as noise.

[Figure 1: see original paper] shows the periodic analysis of sin functions with different data points using the Jurkevich method. To test the reliability of the Jurkevich method for analyzing celestial periods, we took 80-360 data points here, increasing by 20 data points per group, and obtained 15 groups of data points through Jurkevich method analysis, resulting in curves and analysis results of $\langle MATH_1 \rangle$.

[Figure 2: see original paper] shows the periodic analysis of sin functions with different data points using the Date-Compensated Discrete Fourier Transform method (DCDFT). To test the reliability of DCDFT for analyzing celestial periods, we took 80-360 data points, increasing by 20 data points per group, and used DCDFT to analyze 15 groups of data points, obtaining DCDFT-Frequency curves and analysis results of $\langle MATH_2 \rangle$.

[Figure 3: see original paper] shows the periodic analysis of sin functions with different data points using the Discrete Correlation Function method (DCF). To test the reliability of DCF for analyzing celestial periods, we took 80-360 data points, increasing by 20 data points per group, and used DCF to analyze 15 groups of data points, obtaining DCF-Delay curves and analysis results of $\langle MATH_3 \rangle$.

[Figure 4: see original paper] shows the periodic analysis of sin functions with different data points using the Power Spectral Density method (PSD). To test the reliability of PSD for analyzing celestial periods, we took 80-360 data points, increasing by 20 data points per group, and used PSD to analyze 15 groups of data points, obtaining Power-Frequency curves and analysis results of $\langle MATH_4 \rangle$.

[Figure 5: see original paper] shows the minimum data points required for periodic analysis of the sin function using DCDFT, DCF, Jurkevich, and PSD methods. Analysis of [Figure 5: see original paper] reveals that the minimum data points required for period detection by the Jurkevich, DCDFT, DCF, and PSD methods are 50, 120, 100, and 60, respectively. After obtaining the shortest continuous data sampling, the Jurkevich method is most effective.

2.2 Determining Optimal Parameters for the Jurkevich Method Using Simulated Data

When using the Jurkevich method for period analysis, we simultaneously take 360, 720, and 1500 data points, with grouping numbers $m = $, corresponding to . Using a step size of rad1.0 and trial periods, we can calculate different $\langle MATH_5 \rangle$.

[Figure 6: see original paper] shows the $V_m - 2$ curve of sin, with grouping numbers 1-6 on the left and 7-12 on the right. As shown in [Figure 6: see original paper], when taking 360 data points, with grouping numbers increasing from 1 to 6, the requirement is that the number of periods in non-uniform data samples should not be less than 6 [7], so periodicity can be ignored. When grouping numbers increase from 7 to 12, possible double periods appear in groups 10, 11, and 12, which may be repetitions of the first period. Even with minimal differences showing pseudo-periods, the best periodicity is clearly in group 9.

[Figure 7: see original paper] shows the $V_m - 2$ curve of sin, with grouping numbers 1-6 on the left and 7-12 on the right. When taking 720 data points, as shown in [Figure 7: see original paper], with grouping numbers increasing from 1 to 6, there is similarly no periodicity. When grouping numbers increase from 7 to 12, double periods appear in groups 10 and 12, which may be repetitions of the first period, while an obvious pseudo-period appears in group 11. Therefore, the best periodic analysis result is in group 9.

[Figure 8: see original paper] shows the $V_m - 2$ curve of sin, with grouping numbers 1-6 on the left and 7-12 on the right. As shown in [Figure 8: see original paper], when taking 1500 data points, with grouping numbers increasing from 1 to 6, periodicity is not obvious and can be ignored. When grouping numbers increase from 7 to 12, double periods appear in groups 8, 10, and 12, while the periodicity in group 11 is clearly less pronounced than in group 9. Overall, the best periodic analysis result is in group 9. The following will use two sources with multiple observational data to verify this result.

3 Analysis of Light Variation Periods for 3C 279 and 3C 454.3

3.1 Data Points and Light Variation Period of 3C 279

The B-band data points for 3C 279 and 3C 454.3 studied in this paper were mainly obtained from the website (http://www.astro.yale.edu/smarts/glast/home.php). From 2008 to the present, 661 B-band data points for quasar 3C 279 can be found from the above website. The B-band light curve is shown in [Figure 9: see original paper], showing several major outbursts, with the brightest magnitude $m_B = 14.9$ mag. The light curve contains more than 6 periods, which satisfies the necessary conditions for confirming the existence of long periods in the Jurkevich method [13].

[Figure 10: see original paper] shows the Jurkevich method analysis of the light variation period of quasar 3C 279 in the B band. Using the Jurkevich method with grouping number $m = 9$, [Figure 10: see original paper] yields the relationship curve between nP and $2mV$ for the B-band light variation period of 3C 279, which contains many minimum values. To effectively determine the authenticity of periods, reference [14] provides good criteria: when $2mV = $, then $f = 0$, indicating no periodicity in the sample data. If $2mV = $, then $f = 1$, at which point the maximum period in the sample can be identified. Further analysis shows that V_m is a normalized value. When $2mV = $, the sample data shows very strong periodicity. In the normalized graph, selecting the minimum depth and noise in the "smooth"part for experiments, if the relative minimum burst value in the "smooth" part is 10 times larger than the "smooth" part, the corresponding periodicity can be further discussed. The number of groups m is very important in calculations. The more groups m is divided into, the higher the sensitivity, but a small number of data points in each group will produce greater noise. When more groups m are selected, it also increases the computational load of V_m. Conversely, it may not be possible to find the period. Through analysis of the $V_m - 2$ graph for the 3C 279 celestial object, we can easily find the $2mV$ and P values and test them using the criteria from reference [14], with results shown in .

shows the period analysis table for the 3C 279 object. Through analysis of , we find that the minimum value for the 3C 279 object is , with $f = 1.8401$, and the possible light variation period is $\langle MATH_6 \rangle$. Using the Jurkevich method with grouping number $m = 9$, we analyzed the B-band light variation data of 3C 279, with [Figure 10: see original paper] showing the B-band analysis results. From , it can be seen that under the condition of satisfying the criterion f [14], there exists a simple multiple relationship between P2, P3, and P1 in the B band of 3C 279: $P2 \approx 2P1$, $P3 \approx 3P1$, indicating that there may be an astronomical frequency multiplication relationship between them.

3.2 Data Points and Light Variation Period of 3C 454.3

3C 454.3 (PKS 2251+158, OY091) is a relatively bright, violently variable quasar with obvious correlations in optical and radio bands [15]. The data in this paper spans from 2008 to the present, with approximately 765 data points, yielding the historical light curve shown in [Figure 11: see original paper].

[Figure 11: see original paper] shows the light curve of quasar 3C 454.3 in the B band. Using the Jurkevich method with grouping number $m = 9$, we analyzed the B-band light variation data of 3C 454.3, with [Figure 12: see original paper] showing the B-band analysis results. From [Figure 12: see original paper], it can be seen that there are three obvious minimum values in the B band of 3C 454.3, indicating three possible periods: $P1 = 457$d, $P2 = 891$d, and $P3 = 1320$d. However, according to Kidger et al., to determine a period, the time span of the data sample must exceed six times the period [16]. Our data sample spans approximately 2500 days, which is less than six times $P2$ and $P3$, so $P2$ and $P3$ must be excluded and require more observational data for determination. But very interestingly, there exists a simple multiple relationship between $P2$, $P3$, and $P1$: $P2 \approx 2P1$, $P3 \approx 3P1$, indicating that there may be an astronomical frequency multiplication relationship between them. This suggests that there may be a true light variation period of $P1 = 457$d in the B-band light curve of 3C 454.3.

Similarly, through analysis of the $V_m - 2$ graph for the 3C 454.3 celestial object, we can easily find the $2mV$ and P values and test them using the criteria from reference [14], with results shown in .

shows the period analysis table for the 3C 454.3 object. Through analysis of , we find that when the minimum value of $2mV$ for the 3C 454.3 object is 0.5505, with $f = 0.817$, the corresponding possible light variation period is $P = 457$d, which satisfies both the analysis results of [Figure 12: see original paper] and the f criterion [14].

4 Discussion and Conclusions

This paper uses a sine function with period $\pi 2$ as a test signal to verify the reliability of four research methods. In the experiment, the unit is radians, with rad0 as the starting point and a step size of rad1.0. The analysis results obtained are: Date-Compensated Discrete Fourier Transform (DCDFT) method yields $\langle MATH_7 \rangle$; Discrete Correlation Function (DCF) method yields $\langle MATH_8 \rangle$; Power Spectral Density (PSD) method yields $\langle MATH_9 \rangle$. The analysis shows that the minimum data points required for period detection by the Jurkevich, DCDFT, DCF, and PSD methods are 50, 120, 100, and 60, respectively. After obtaining the shortest continuous data sampling, the Jurkevich method is most effective. The Jurkevich method' s analysis results are the most accurate and reliable among the four methods studied, and the calculation method is simple and practical.

In practical application, it was found that the Jurkevich method depends heavily on the grouping number m. Larger m values yield better analysis results but produce greater noise; conversely, periods may not be found [17]. If the data distribution is non-uniform, it leads to large deviations in inter-group data distribution, thereby affecting the actual period determination. According to [Figure 6: see original paper], [Figure 7: see original paper], and [Figure 8: see original paper], for different m values, the inclination degree of the nP versus $2mV$ relationship curve differs, with larger periods tilting downward. Combining this with the f test formula shows that smaller mV yields larger f values, making it easier to satisfy the f test. Thus, the f test depends on the magnitude of m. Using simulated data, we found that the optimal grouping number for the Jurkevich method is $m = 9$. Finally, using the Jurkevich method with $m = 9$, we analyzed the light variation periods of 3C 279 and 3C 454.3. We obtained a possible light variation period of for 3C 279. Literature [18] analyzed a possible period of $130 ±$ for 3C 279, which is about 8 times our result. The possible light variation period for 3C 454.3 is 457d. Literature [19] analyzed a possible period of 12.39 years for 3C 454.3, which is about 10 times our obtained period.

Acknowledgments: We thank the editors and reviewers for their comments on our manuscript "Study on long-period variability analysis methods of quasars." These comments were very helpful for revising and improving the quality of our paper. We have carefully revised the paper according to the reviewers' suggestions.

References

[1] Xie Zhaohua. Study of jet-disk symbiosis and accretion disk viscosity parameter of Blazars[D]. Yunnan University, 2012.

[2] Wang Wenguang. Analysis of Optical Variability and Color Index of Blazar[D]. Journal of Yunnan Normal University, 2017.

[3] Fan J H, Adam G, Xie G Z, et al. Correlation between the gamma-ray and the radio emissions[J]. Astronomy & Astrophysics, 1998, 338(1):27-30.

[4] Fan J H, Kurtanidze O, Liu Y, et al. Optical Monitoring of Two Brightest Nearby Quasars, PHL 1811 and 3C 273[J]. Astrophysical Journal Supplement, 2014, 213(2):26.

[5] Rieger F M, Mannheim K. Implications of a possible 23 day periodicity for binary black hole models in Mkn~501[J]. Astronomy & Astrophysics, 2012, 359(3):948-952.

[6] Webb J R, Smith A G, Leacock R J, et al. Optical observations of 22 violently variable extragalactic sources - 1968-1986[J]. Astronomical Journal, 1988, 95(2):374-397.

[7] Jurkevich J.N. A METHOD OF COMPUTING PERIODS OF CYCLIC PHENOMENA[J]. Applied Surface Science, 1971, 13:154-167.

[8] Yu Lian, Zhang Xiong, Wang Wenguang, Luo Shuangling. The Variability Analysis of PKS 0735+178[J]. Astronomical research and technology, 2018, 15(01):10-16.

[9] Ferraz-mello S. Estimation of Periods from Unequally Spaced Observations[J]. Astronomical Journal, 1981, 86(86):619.

[10] Luo Shuangling, Zhang Xiong, Wang Wenguang. The Optical Variability Periodicity Analysis of PKS 1510-089 Based on Date Compensated Discrete Fourier Transform[J]. Journal of Yunnan Normal University: Natural Sciences Edition, 2016, 36(5): 1-4.

[11] Fan J H, Lin R G. The variability analysis of PKS 2155-304[J]. Astronomy & Astrophysics, 2000, 355(3):880-884.

[12] Zhang Rongzhu, Cai Bangwei, Yang Chunlin, Xu Qiao, Gu Yuanyuan. Numerical Method of the Power Spectral Density[J]. High Power Laser and Particle Beams, 2000(06): 661-664.

[13] Zhang Xiong, Xie Guangzhong, Bai Jinming. Computing the Period of Light Variability in BL Lac Objects using the Jurkevich Method[J]. ACTA ASTROPHYSICA SINICA, 1998(03): 33-41.

[14] Kidger, M.R., Takalo, L., & Sillanp, A. A new analysis of the 11-year period in OJ287: confirmation of its existence[J]. Astronomy and Astrophysics, 1992, 264:32-361.

[15] PYATUNINA T B, KUDRYAVTSEVA N A, GABUZDA D C, et al. Frequency-dependent time delays for strong outbursts in selected Blazars from the Metsahovi and UMRAO monitoring data bases- [J]. Monthly Notices of the Royal Astronomical Society, 2007, 381(2): 797-808.

[16] Heidt J, Wagner S J. Statistics of optical intraday variability in a complete sample of radio-selected BL Lac objects[J]. Astronomy and Astrophysics, 1996, 305:42-52.

[17] Liu KaiBo, Yang Dongmo, Hou Dedong, Zhang Xiong, Ding Liang, Dong Futong. Computing the Period of Light Variability in Seyfert Galaxy 3C 120 Using the Jurkevich Method[J]. Journal of Yunnan Normal University: Natural Sciences Edition, 2009, 29(05):1-5+16.

[18] Li Huaizhen. Study on the light and energy spectrum distribution of Blazar celestial bodies[D]. Yunnan University, 2011.

[19] Su Chengyue. Long-term optical photoperiod analysis of quasar PKS 2251+158[J]. Research in Astronomy and Astrophysics, 2000, 20(1):11-16.

Note: Figure translations are in progress. See original paper for figures.