Jets at all Scales
Proceedings IAU Symposium No. 275, 2011
G. E. Romero, R. A. Sunyaev & T. Belloni, eds.

doi:10.1017/S1743921310015826

Variability studies in blazar jets with SF analysis: caveats and problems

Dimitrios Emmanoulopoulos[1], Ian M. McHardy[1] and Phil Uttley[1]

[1]School of Physics & Astronomy, University of Southampton
SO17 1BJ, Southampton, United Kingdom
email: D.Emmanoulopoulos@soton.ac.uk

Abstract. Blazars are radio-loud active galactic nuclei (AGN) dominated by relativistic jets seen at small angles to the line-of-sight. They exhibit dramatic flux variations across the electromagnetic spectrum. The fastest variations are observed in the X-ray and γ-ray bands on time-scales of hours or even minutes. Currently, a substantial part of the blazar literature has been based on the study of these temporal variations through the use of structure function (SF) analysis, the results of which are believed to put great constrains on the jet-physics.

The SF is often invoked in the framework of shot-noise models to determine the temporal properties of individual shots within the jet as well as their geometrical sizes. We argue, that for blazar variability studies, the SF-results are sometimes erroneously interpreted leading to misconceptions about the actual source properties. Based on extensive simulations we caution that spurious breaks will appear in the SFs of almost all light-curves, even though these light-curves may contain no intrinsic characteristic time-scale.

Finally, it is also commonly thought that SFs are immune to the sampling problems, such as data-gaps, which affects the estimators of the underlying power spectra density function such as the periodogram. However, we show that SFs are also troubled by gaps which can induce artefacts.

Keywords. (galaxies:) BL Lacertae objects: general, galaxies: jets, X-rays: general, methods: statistical

1. Intordudction

One of the most extensively used tools in the field of blazar variability is the structure function (SF) (e.g. Hughes *et al.* 1992) which measures the mean value of the flux-variance for measurements, $x(t)$, that are separated by a given time interval, τ, where $SF(\tau) = \left\langle [x(t) - x(t+\tau)]^2 \right\rangle$. The SF is commonly characterized in terms of its slope β, where $SF(\tau) \propto \tau^{\beta}$.

We will show, through extensive simulations, that much of the existing literature on blazar SFs tends to misinterpret observed SF-characteristics, such as breaks, as being real or physically meaningful. Often the SF-breaks are either artefacts intrinsic to SFs, or subject to much greater statistical variation than inferred from the commonly-used fitting procedures

We create artificial light-curves, lacking any sort of characteristic time-scales, and we study the SF-behaviour of MRK 501, derived from both the thoroughly studied *ASCA* data-set Tanihata *et al.* (2001) (TAN01), and the long-look light-curve of *All-Sky Monitor* (*ASM*) (onboard (*Rossi X-ray Timing Explorer*)), *RXTE*). Finally, we test the statistical robustness of the most commonly used fitting-procedures which are employed in order to derive astrophysically interesting quantities from the SF, as well as the sensitivity of SF to the presence of data-gaps. A complete description of our simulations and our results can be found in Emmanoulopoulos *et al.* (2010) (EM10).

2. SF and data-length: MRK 501

In this section we examine the SF-behaviour of MRK 501 by employing the *ASCA* data-set of TAN01 in 2–10 keV (∼10 days), with a sample interval of 5678.3 sec, and the ASM data-set (∼2000 days), in bins of 15 days.

We simulate stationary light-curves based on the procedure described by Timmer & Koenig (1995), which uses as an input the power spectral density (PSD) function of the observed light-curve. The output of the method is an ensemble of stochastic time-series produced having the same statistical properties with the observed data-set. Moreover, this methodology allows us to take into account the *red-noise leak* and *aliasing* effects.

Initially, we estimate the periodogram of the ASCA data-set which can be very well fitted by a simple featureless power-law with index $\alpha = -1.80 \pm 0.09$ ($\chi^2 = 3.80$ for 6 degrees of freedom with a null hypothesis probability of 0.70). Next we produce 2000 artificial light-curves having the same power-law PSD of index of $\alpha = -1.80$ and the same length as the studied *ASCA* data-set. For each artificial light-curve we estimate the SF and we localize the position of the break by producing an interpolated version of each SF. The distribution of the position of the SF-breaks is shown in the left panel of Fig.1 (grey area) having a mean value of 0.92 and and a standard deviation of 0.09 days. These simulations show us clearly that stochastic data-sets having the same length as the *ASCA* data-set of MRK 501 and the same featureless PSD, exhibit breaks in the SF around a day. Thus, the apparent break seen in the SF for MRK 501 from TAN01 should not be associated with any sort of physically meaningful time-scale.

By extending our simulations to even longer time-scales (2000 days long), we can check the effects of the data-length in the SF-estimates. From a total of 2000 long-term light-curves 1893 exhibit SF-breaks whose distribution is well represented by a Gaussian distribution having a mean 399.35 days and a standard deviation of 74.73 days (Fig.2, right panel in EM10).

In order to compare our predicted break with real data we employ the long-term *RXTE-ASM* light-curve of MRK 501 (Emmanoulopoulos 2008) and we estimate its SF. The SF-break occurs around 402 days (Fig.3, right panel in EM10), something which is absolutely in accordance with the results from our simulated light-curves of the same length which do not include any sort of characteristic time-scale.

The aforementioned example reflects in a clear way that the SF deals only with the properties of the observed light-curve, ignoring the properties of the true underlying variability process taking place within the jet.

3. SF and gappy data-sets

The SF method is considered by several researchers (e.g. Kataoka *et al.* 2001, Kataoka *et al.* 2002, Zhang *et al.* 2002) to be the ideal method of studying the time properties of gappy data-sets as it is believed to be less distorted by gaps than frequency-domain methods.

We produce an ensemble of 2000 artificially produced light-curves, having PSD slope of −1.5 and being 2000 t.u. long, with a given gappy pattern (EM10). Then we apply the bootstrap method (e.g. Czerny *et al.* 2003) 1000 times to each one of the aforementioned light-curves, we estimate for each SF bin i the quantity $|\log_{10}[SF^i_{\mathrm{gappy}}(\tau)] - \log_{10}[SF^i_{\mathrm{conti}}(\tau)]| \cdot err_{\mathrm{boot},i}(\tau)^{-1}$, and finally we estimate its mean value. As we can see

https://doi.org/10.1017/S1743921310015826 Published online by Cambridge University Press

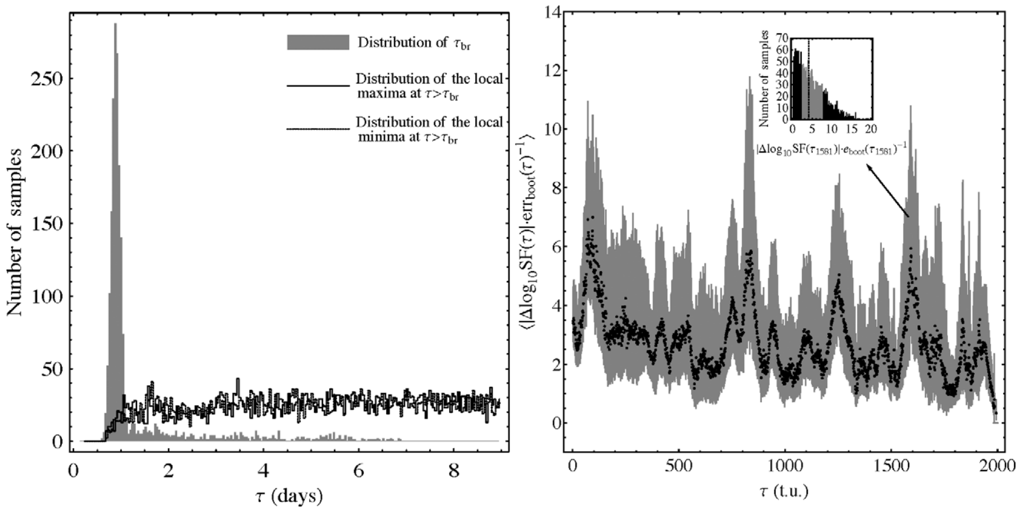

Figure 1. [Left panel] The grey area depicts the distribution of the SF-breaks coming from the 2000 simulated light-curves (the histogram-bins have a length of 0.04 days). The solid and dashed lines represent the distribution of the local maxima and local minima for $\tau > \tau_{\rm br}$ mapping the positions of the wiggling features. [Right panel] The mean difference between gappy and continuous SFs coming from an ensemble of 2000 artificial light-curves.

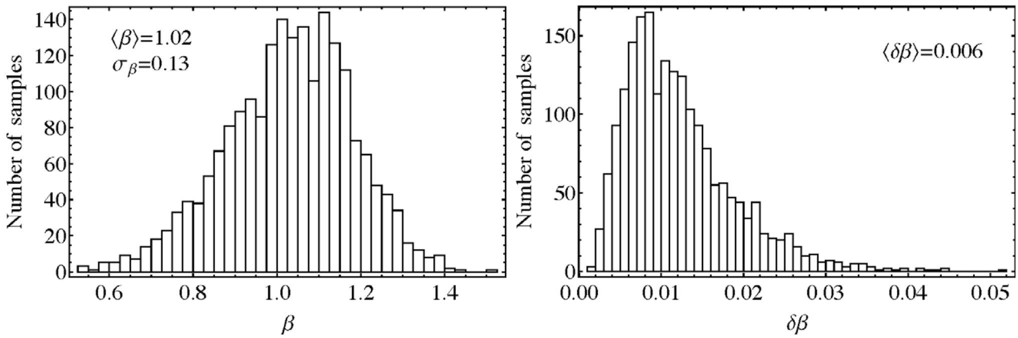

Figure 2. [Left panel] The distribution of the SF slopes β has a mean of 1.02 and a standard deviation 0.13 (the histogram-bins have a length of 0.025). [Right panel] The distribution of the errors coming from the fit $\delta\beta$, having a mean of 0.006 (the histogram-bins have a length of 0.001).

from the right panel of Fig.1, the various estimates differ significantly from unity, having a mean value of 2.79 ± 0.91. That means that data-gaps do affect the SF-results in an erratic way by introducing systematic deviations which depend on the light-curve realization. Moreover, the bootstrap method does not yield statistically meaningful errors that reflect the true deviations between the continuous and the gappy SFs.

4. SF and fitting procedures

One of the major problems that affects the results of the SF-analysis is that the various estimates, $SF(\tau_i)$, are not statistically independent of each other. This problem affects severely the fitting routines e.g. least-squares, maximum likelihood, that are commonly used in the published blazar-SF-literature to derive the SF-breaks and the -slopes, yielding very small estimates of uncertainty in the fitting parameters (EM10).

https://doi.org/10.1017/S1743921310015826 Published online by Cambridge University Press

Having produced 2000 light-curves, 500 time-units (t.u.) long from a power-law PSD having an index -2, we calculate for each one of them the logarithm of the SF-estimates. We attribute to every estimate a standard error based on the error of the sample mean for each time bin, which is one of the most commonly used methods (e.g. Zhang *et al.* 2002, Czerny *et al.* 2003). From Fig.2 we can see that the errors $\delta\beta$ derived from the fit for the individual SF-slopes β (right panel) are very small in comparison to the actual scatter of the fitted β (left panel).

Since the SF-slope β is associated with the nature of the variability process within the jet (e.g. red-noise, white-noise), different realizations of the same variability process may appear having significantly different SF-slopes due to error underestimation.

5. Conclusions

Our results can be summarized as follows:

• Strong SF-breaks frequently occur in data-sets lacking any sort of characteristic time-scales. The position of these physically uninteresting breaks depends on the length of the observations and the shape of the underlying PSD.

• Data-gaps affect severely the SF-estimates in an unpredictable way, introducing systematic deviations. The bootstrap method can not yield statistically meaningful errors that depict the true deviations between the gappy and the continuous SFs.

• Non-independence and non-Gaussianity make impossible the estimation of a meaningful goodness-of-fit from based on normal fitting procedures. We see, for example, that the derived uncertainties on the SF quantities i.e. positions of breaks and slopes, are always much smaller than the actual scatter of these variables during multiple realizations of the same variability process.

Acknowledgements

DE and IMM acknowledge the Science and Technology Facilities Council (STFC) for support under grant ST/G003084/1. PU is supported by an STFC Advanced Fellowship. This research has made use of NASA's Astrophysics Data System Bibliographic Services.

References

Czerny B., Doroshenko V. T., Nikołajuk M., Schwarzenberg-Czerny A., Loska Z., & Madejski G. 2003, *MNRAS*, 342, 1222

Emmanoulopoulos, D. 2008, *Blazar Variability across the Electromagnetic Spectrum*, *PoS*, (BLAZARS2008)038, Palaiseau, France

Emmanoulopoulos, D., McHardy, I. M., & Uttley, P. 2010, *MNRAS*, 404, 931

Hughes P. A., Aller H. D., & Aller M. F. 1992, *ApJ*, 396, 469

Kataoka, J., Takahashi, T., Wagner, S. J., Iyomoto, N., Edwards, P. G., Hayashida, K., Inoue, S., Madejski, G. M., Takahara, F., Tanihata, C., & Kawai, N. 2001, *ApJ*, 560, 659

Kataoka, J., Tanihata, C., Kawai, N., Takahara, F., Takahashi, T., Edwards, P. G., & Makino, F. 2002, *MNRAS*, 336, 932

Tanihata C., Urry C. M., Takahashi T., Kataoka J., Wagner S. J., Madejski G. M., Tashiro M., & Kouda M. 2001, *ApJ*, 563, 569

Timmer J. & Koenig M. 1995, *A&A*, 300, 707

Zhang, Y. H., Treves, A., Celotti, A., Chiappetti, L., Fossati, G., Ghisellini, G., Maraschi, L., Pian, E., Tagliaferri, G., & Tavecchio, F. 2002, *ApJ*, 572, 762

Discussion

Kawai: What is your suggestion for estimating the characteristic time scale from data with irregular gaps?

Emmanoulopoulos: The best way is to use the Fourier domain, i.e.m power spectral density, even in the case of gappy irregular data sets. Through simulations you can take into account aliasing and sampling effects in a robust and statistically correct way (e.g. Uffley *et al.* 2009)

Meier: So, should we then conclude that blazars have no preferred time scale, i.e., that their PSD is simply a power law?

Emmanoulopoulos: Yes, the PSD of blazars (up to now) is a simple power law, having no breaks or other features.