

# On the time delay evolution of five Active Galactic Nuclei

A. Kovačević, L. Č. Popović, A. I. Shapovalova, D. Ilić, A. N. Burenkov and V. H. Chavushyan

**Abstract** Here we investigate light curves of the continuum and emission lines of five type 1 active galactic nuclei (AGN) from our monitoring campaign, to test time-evolution of their time delays. Using both modeled and observed AGN light curves we apply Gaussian-kernel based estimator to capture variation of local patterns of their time evolving delays. The largest variations of time delays of all objects occur in the period when continuum or emission lines luminosity is the highest. However, Gaussian kernel based method shows instability in the case of NGC 5548, 3C 390.3, E1821+643 and NGC 4051 possible due to numerical

---

A. Kovačević, D. Ilić

Department of astronomy, Faculty of mathematics, University of Belgrade, Studentski trg 16, 11000 Belgrade, Serbia e-mail: andjelka@matf.bg.ac.rs, dilic@matf.bg.ac.rs

L. Č. Popović

Astronomical Observatory Belgrade, Volgina 7, 11060 Belgrade, Serbia, e-mail: lpopovic@aob.rs

A. I. Shapovalova, A. N. Burenkov

Special Astrophysical Observatory of the Russian AS, Nizhnij Arkhyz, Karachaev - Cherkesia 369167, Russia, e-mail: ashap@sao.ru, banbur@gmail.com

V. H. Chavushyan

Instituto Nacional de Astrofísica, Óptica y Electrónica, Apartado Postal 51y 216, 72000 Puebla, México e-mail: vahram@inaoep.mx

discrepancies between Damped Random Walk (DRW) time scale of light curves and sliding time windows of the method. The temporal variations of time lags of Arp 102B can correspond to the real nature of the time lag evolution.

Keywords: galaxies: active ;galaxies:individual: 3C 390.3, Arp 102B, NGC 5548, NGC 4051 and E1821+643;Methods: statistical; line profiles

## 1 Introduction

The timescales of active galactic nuclei (AGN) optical variability range from weeks up to years, which causes data analysis problems that are not usually presented in the case of the fast variable objects. The large number of observations and good sampling covering the full range of frequencies has almost never been possible to achieve in AGN observation. AGN time series are obtained from monitoring campaigns consisting of regular observations over long period, or continuous observations set on much shorter time span. AGNs have strong emissions lines (e.g. the Balmer line series), excited by photoionisation from the continuum emission. These lines respond to the variations of the ionizing continuum, and time delays between continuum and line variations can be related to the distance (geometry) between continuum and line emitting regions. Since AGN do not vary periodically, their delays are seen as echoes of the variations of the continuum. Also, the line response to the continuum can be non-linear, non-stationary at some level. For example, in case of NGC 5548 (see Cackett and Horne(2005), and reference therein) the H $\beta$  lag relative to the continuum is determined, on a year by year basis using CC, and it was found that lag increases with the increase of the mean continuum flux. Also, various studies shown that emission lines have a nonlinear response to continuum variations (Pogge and Peterson (1992), Deitrich and Kollatschny (1995)).

Rehfeld and Kurths (2014) investigated similarity estimators suitable for the quantitative measure of dependencies in irregular and/or non-linear time series. They introduced Gaussian-kernel-based CC (captures linear relationship) and Gaussian mutual information (MI) measure (captures non linear relationship), and found that interpolation to regular spacing of the observation times results in worse estimates, while their adapted estimators are more efficient in the presence of sampling time irregularity. According to their approach it is possible to calculate windowed cross similarity between signals and to see how the time delay is changing over the time span covered by observations.

Mostly in CC analysis of AGNs, the underlying hypothesis is that estimated quantities do not vary significantly for the duration of the cross-correlated signals. We investigated CC and time delays of continuum and emission line curves of 4 type 1 AGNs: Arp 102B, 3C 390.3, NGC 5548, and NGC 4051 with assumption that those parameters do not vary with time in our previous work (see Kovačević et al. 2014). Our analysis is focused here on the detection of time evolution of linear (measured by CC) and non linear relationships (measured by (MI)) of time series of our data set enlarged with time series of E1821+643. Some geometrical characteristics of broad line region (BLR) can be extracted from the time delays , so any detection of variation of time delays over monitored period can contain information about its geometry variations. The structure of the paper is organized as follows. In the section Data a short description of used data samples is given. In the section Method Gaussian kernel based CC and MI method are described. In the section Results and Discussion the maps of time delay variations during observed periods are displayed and discussed their implications. In the section Conclusion our main findings are summarized.

## 2 Data

We analyzed the integral fluxes of the broad H $\alpha$  and H $\beta$  emission lines and the fluxes of the continuum of our objects. Spectra of Arp 102B, 3C 390.3, NGC 5548, and E1821+643 are from our monitoring campaign Shapovalova et al. (2004), Shapovalova et al. (2010), Shapovalova et al. (2013), Shapovalova et al. (2015). As for NGC 4051 the data were taken from Denney et al. (2006). Detailed data and discussion of observations and data reduction could be found in papers mentioned above.

The sampling pattern (Pessah 2005) is characterized by vector  $SP = (T, T_S, T_B)$ , which components stand for the duration of the observation campaign ( $T$ ), the sampling interval ( $T_S$ ), and the binning interval ( $T_B$ ). Since our sets of data are obtained from optical monitoring campaigns, their  $T_S = T_S(T)$  is discrete variable over the time-period of observation. We calculated the rms amplitude  $\sigma_{rms}^2$  (Pessah 2005) associated with each light curve flux L as

$$\sigma_{rms}^2 = \frac{1}{P-1} \sum_{n=0}^{P-1} |L - \bar{L}|^2 \quad (1)$$

where  $P = \frac{T}{T_S}$ ,  $\bar{L}$  and  $\bar{T}_S$  are the average values corresponding to the particular time series and its  $T_S$  respectively. Comparison of  $\sigma_{rms}^2$  for all objects as function of the ratio between the observing and the sampling times is given in Table 1.

An example of nonlinearity of our time series is given on the upper panel in the Fig. 1. Simple visual inspection of the 6-year long NGC 5548 light curves shows that in the lowest state (in 2002) of the H $\beta$  line is more concave than the same feature in the continuum, while in the highest state (1998), the H $\beta$  peak is not as pronounced as the continuum's peak. Since the light travel time delay blur out the emission line response, this blurring effect should happen to both the peak and the concave feature simultaneously. The same pattern occurred in 1989-2001 during

| Object    | Light curve | $\frac{T}{T_S}$ | $\sigma_{rms}^2$ |
|-----------|-------------|-----------------|------------------|
|           | cnt         | 115             | 2.30             |
| Arp 102B  | H $\alpha$  | 89              | 40.38            |
|           | H $\beta$   | 117             | 2.06             |
|           | cnt         | 127             | 1.16             |
| 3C 390.3  | H $\alpha$  | 47              | 11.18            |
|           | H $\beta$   | 128             | 0.87             |
|           | cnt         | 80              | 12.37            |
| NGC 5548  | H $\alpha$  | 55              | 81.02            |
|           | H $\beta$   | 83              | 6.11             |
| NGC 4051  | cnt         | 234             | 0.19             |
|           | H $\beta$   | 107             | 0.13             |
| E1821+643 | cnt         | 126             | 1.89             |
|           | H $\beta$   | 126             | 0.52             |

Table 1: Sampling pattern ( $\frac{T}{T_S}$ ) and the rms amplitude ( $\sigma_{rms}^2$ ) for continuum (cnt),  $H\alpha$  and  $H\beta$  observed light curves.

AGN Watch monitoring campaign of NGC 5548 (see details in Cackett and Horne (2005)).

## 2.1 Method

From general point of view, the CC (or similarity) allows discovering groups of objects with similar behavior and, potential anomalies which may be revealed by change in CC.

Specifically the time lags between correlated time series of unalike but related origin reveals the causalities (linear connections) within the system. Namely, the CC coefficients are calculated based on hypothesis that pairs of values from each series are subjected to simple linear relationship. In order to reconstruct linear

dependencies of hidden processes from the observations, the cross power spectra and CC are used due to Fourier transform connection between them (Chatfield (2004)). However, the irregularity of time series sampling inhibits application of standard methods of similarity (cross power and CC) measures. There are several approaches to overcome this problems, classified generally as (Broersen and de Waele (2000)): (i) direct transform methods, (ii) slotting techniques, (iii) model-based estimators, and (iv) time series reconstruction methods. The Lomb-Scargle periodogram (Lomb (1976), Scargle (1981), Scargle (1982)) is a well-known type (i) method that computes a least squares fit of sine curves to the data. Lomb-Scargle is powerful in finding narrow peaks with a low noise level. However, it loses its strength with smaller peaks and more noise.

Standard slotted CC estimation discretizes the distance between two observations to slots of width  $\Delta$ . The product of two irregular observations contributes to the slotted autocorrelation at a certain lag  $k\Delta$ , if their distance is within the range  $(k \pm 0.5)\Delta$  (Mayo (1993), Edelson and Krolik (1988)). In such a way, this technique interpolate in time delay domain (contrary to the interpolation in the time domain for one or both members of the correlation (see Gaskell and Sparke (1986), Korista and Goad (1991)). The disadvantage of this technique is the correlation function estimates are not necessarily positive semidefinite and the spectra computed from their Fourier transform can show negative power. So some kind of post-processing techniques are introduced to avoid this problem (Stoica et al. (2008), Babu et al. (2009)).

Model-based estimators fit a model to the time series and then try through some kind of transformation to determine the time lag between time series, such example is SPEAR method (Zu et al. (2011)). The fourth group of estimators consists of different kind of interpolation methods in time domain of time series. The case of the standard linear interpolation of irregular observations, onto an even sampling grid, adds bias towards low frequencies in power spectral density (PSD) estimation.

Application of cubic spline interpolation can lead to spurious peaks, so the variance of resampled signal is too high. As opposed to cubic spline, linear interpolation is a robust resampling method, it does not add spurious peaks, but the variance of resampled signal is too low. This arises from the fact that linear interpolation is weighted average of two irregular observations, so the resampled signal will have lower variance than original one.

Rehfeld et al. (2011) probed different CC estimators for irregular time series and found that a Gaussian-kernel-based estimator performed best, which has the form introduced in Bjoernstad and Falck (2001) :

$$ker(k - \Delta t_{ij}^{xy}) = ker(d) = \frac{1}{\sqrt{2\pi h}} \exp^{\frac{-|d|^2}{2h^2}} \quad (2)$$

where  $\Delta t_{ij}^{xy}$  is the inter-observation time, and k denotes the k-th lag. The CC function can be written as

$$CC(k\Delta\tau) = \frac{\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} x_i y_j ker}{\sum_{i=1}^N \sum_{j=1}^N ker} \quad (3)$$

where  $x_i, y_j$  and  $N_x, N_y$  are time series values and their lengths respectively,  $k\Delta\tau$  is k-th time lag. Gaussian window has three desirable properties:(i) optimal resolution. Only the Gaussian window minimizes the resolution product of its width and width of its Fourier transform. (ii) Isotropic (circularly symmetric). Isotropy implies no directional bias in multi-dimensional correlations. (iii) Separable in N dimensions. It implies computational efficiency. (iv) the product of two Gaussians with equal widths is Gaussian.

In our analysis we are interested to track any possible time-evolving presence in our data, by means of time-evolving correlation function. Rehfeld method allows splitting signals into overlapping segments and calculating time-evolving Gaussian kernel CC.

On the other hand we know that the covariance functions of the AGN light curves are close to exponential kernels, i.e, the damped random walks (DRW), on scales

beyond a week, and probably have a break on smaller scales based on Mushotzky et al. (2011) analysis of the Kepler mission light curves. It has been confirmed by Kelly et al. (2009), Kozlowski et al. (2010), Macleod et al. (2010), Zu et al. (2013), Andre et al. (2013).

Also, we are interested to measure any possible non-linear correlations among our time series (the case of NGC 5548 is known as non-linear, see Fig. 1). *MI* can measure both linear and nonlinear relationship between time series.

*MI* strictly determines (in units of bits) how much information the value of one variable reveals about the value of another. This is of great importance in information theory (Cover and Thomas (1991)).

Perhaps less well known, is applicability of *MI* on the statistical analysis of data. It is not strange since the mathematical theory of information transmission established by Shannon (1948) represents the culmination of statistical research. Namely the ideas from statistics become applicable to a strikingly wide range of problems. To get an intuitive sense of *MI* notion, let us consider an example of two *independent* random variables  $X$ ,  $Y$ . Since they are independent, their joint distribution  $p(X, Y)$  is identical to the product of marginal distributions  $p(X)$  and  $p(Y)$ . So one could investigate the degree of independence between  $X$  and  $Y$  by computing probabilistic distance between  $p(X)p(Y)$  and  $p(X, Y)$ . This distance in the case of two independent variables is 0. A common probabilistic distance between variables are measured by Kullback-Leibler divergence, and if we apply it to our intuitive example to measure distance between the joint distribution and the marginals of two random variables we will come to the concept of *MI* inevitably. The *MI* ( $I(X, Y)$ ) between the variables  $X = (x_1, \dots, x_n)$  and  $Y = (y_1, \dots, y_m)$  is defined (Shannon (1948), Kolmogorov (1968)) in terms of entropy ( $H(X) = -\sum_{i=1}^n p(x_i) \log p(x_i)$ ) as

$$I(X, Y) = H(X) + H(Y) - H(X, Y) \geq 0 \quad (4)$$

or in terms of their joint probability distribution  $p(X,Y)$  as

$$I(X,Y) = \int dx dy p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)} \quad (5)$$

MI definition assures that its value is always nonnegative and zero only when  $p(X,Y) = p(X)p(Y)$ . The MI will be greater than zero when X and Y are dependent mutually, regardless of the level of nonlinearity of that relationship. The stronger mutual dependance implies the larger the value of  $I(X,Y)$ . One can say that MI is deeply connected to the statistical problem of detecting dependencies. Namely, from previous equation, it is clearly seen that for the data taken from distribution  $p(X,Y)$ ,  $I(X,Y)$  quantifies the expected log-likelihood ratio of the data with underlying  $p(X,Y)$  as opposed to  $p(X)p(Y)$ . So  $I(X,Y)^{-1}$  is the typical amount of data needed to be collected in order to make double increase in posterior probability of variables dependance relative to the hypothesis that they are independent. On the other hand Neyman-Pearson lemma (Neyman-Pearson (1933)) asserts that all useful information about differentiating between two hypothesis is contained in  $\sum_i \log_2 \frac{p(x_i, y_i)}{p(x_i)p(y_i)}$ , possessing the maximal possible statistical power for such test. Based on this conclusion inferred from Neyman-Pearson lemma, one can say that  $I(X,Y)$  (defined to be exactly the limit of such sum) provides upper constrain on how well can be performance of any dependance test applied on data obtained from distribution  $p(X,Y)$ . Estimation of MI is nontrivial, because of nontrivial estimation of the joint distribution  $p(X,Y)$  from a finite sample of N observation. There are several approaches to overcome this problem, and Rehfeld and Kurts (2014) implemented one of them. Similar to the Gaussian kernel based CC, Rehfeld and Kurths (2014) introduced Gaussian kernel based MI. Algorithmically, the procedure consists of constructing a new bivariate set of observations ( $Y^X$ ) by estimating for each point in time series  $X = (t_i^x, x_i)$  a local Gaussian-kernel-based weighted mean in second time series  $(t_i^y, y_i)$ . The same procedure is repeated for the first time series, obtaining  $X^Y$ . Then

original and reconstructed series are concatenated  $(X \cup X^Y), (Y \cup Y^X)$ . The joint density of X and Y is estimated using standard binning estimators for MI. This procedure also can track time evolving of MI.

### 2.1.1 Results and discussion

There is a large theoretical effort to understand the innermost part of active nucleus parametrized by black hole mass, spin and Eddington ratio. However, this process is not stationary imposing additional problems (Czerny and Hrynewicz (2012)). All our time series (see Kovačević et al. (2014) and reference therein) are very fluctuated, with structural change. Therefore, we firstly examine properties of our data sets by using stationarity test. Time series stationarity is a statistical characteristic of a series mean and variance over time. If both are constant over time, then the series is assumed to be a stationary process (i.e. is not a random walk/has no unit root), otherwise, the series is described as being a non-stationary process (i.e. a random walk/has unit root).

Stationarity of a series is an important phenomenon because it can influence its behavior. For example, the term shock (Brooks (2014)) is used frequently to indicate an unexpected change in the value of a variable (or error). For a stationary series a shock will gradually terminate. That is, the effect of a shock during time  $t$  will have a smaller effect in time  $t + 1$ , a still smaller effect in time  $t + 2$ , etc.

Non-stationarity can raise from deterministic changes like trend or seasonal fluctuations or the stochastic properties of processes. In the case of testing for non-stationarity and stationarity of the time series autoregressive unit root tests are valid if the time series  $y_t$  is well described by an AR(1) with white noise errors. However, many time series posses a complex dynamic structure which cannot be captured by a simple AR(1) model. Said and Dickey (1984) augmented the basic autoregressive unit root test to allow for general ARMA(p,q) models with unknown

orders which is called the Augmented Dickey-Fuller (ADF) test. The ADF test probes the null hypothesis that a time series is a non-stationary against the alternative that it is stationary assuming that the underlying data dynamics has an ARMA structure.

The ADF test estimates the test regression in the form

$$y_t = \mathbf{D}_T + \phi y_{t-1} + \sum_{j=1}^p \psi_j \Delta y_{t-j} + \varepsilon_t \quad (6)$$

where  $\mathbf{D}_T$  is a vector of deterministic terms (constant, trend, etc). Imposing constraints on constant and time trend to be zero corresponds to modeling a damped random walk, while setting only time trend to be zero corresponds to modeling a random walk with a drift. The total of  $p$  lagged difference terms,  $\Delta y_{t-1}$  approximate the ARMA structure of the errors. The value  $p$  is set that the error  $\varepsilon_t$  is serially uncorrelated. The error term is assumed to be homoskedastic. The unit root test is then carried out under the null hypothesis  $\phi = 0$  against the alternative hypothesis of  $\phi < 0$ . The ADF t-statistic ( $ADF_t$ ) and normalized bias statistic ( $ADF_N$ ) are calculated by means of the least squares estimates of equation and are given by

$$\begin{aligned} ADF_t &= t_{\phi=1} = \frac{\check{\phi}-1}{SE(\phi)} \\ ADF_N &= \frac{N(\check{\phi}-1)}{1-\check{\psi}_1-\dots-\check{\psi}_p} \end{aligned} \quad (7)$$

where  $\check{\phi}$  is an estimate of the coefficient  $\phi$ ,  $SE(\phi)$  is its standard error in alternative model,  $\check{\psi}_1 \dots \check{\psi}_p$  are stationary coefficient in alternative model, and  $N$  is the effective sample size.

We employed Augmented Dickey-Fuller (ADF) test implemented in SciKits Statmodels of time series analysis in Python language. We tested our time series with a null hypothesis of non-stationarity against an alternative hypothesis of stationarity around:constant, a constant + trend, a constant+ trend + trend squared, and no constant trend respectively. We summarize in Table 2 outputs of

testing our time series around no constant trend, which is equal to modeling damped random walk, as we mentioned above. When we compare the test statistics and critical value, we have to reject the null if test statistics is smaller than critical value (not absolute value, because that is usually applied to two-tailed test). For example if the test statistics is  $-1.85$  and critical value is  $-3.5$  we will not reject the null hypothesis.

Time series of NGC 5548, 3C390.3, H $\beta$  line of NGC 4051 and continuum (5100 Å) of E1821+643 are clearly non stationary in all tests at all confidence levels. While for other time series the non-stationarity is clearly expressed when testing them without constant trend. We accepted all our light curves as non-stationary, since their ADF test values around no constant trend confirm their non-stationarity. Due to non stationarity CC should evolve with time, which we investigated by techniques of Gaussian based CC and MI. The application of these techniques on our data sets is facilitated by a Matlab toolbox (Rehfeld and Kurths (2014), <http://tocsy.pik-potsdam.de/nest.php>), modified for our purposes. The CC coefficients are estimated for each lag shift given in the input vector. The signals are splitted into overlapping segments, both splitting window ( $W$ ) and overlapping parameter ( $O$ ) must be specified. Within the overlap of the individual pairs, we calculate the matrix with their CC coefficients for each lag as specified by input vector of lags. The matrix contains  $N$  rows ( $N$  is the length of input lag's vector) and  $k = \text{numpy.floor}(\text{ceil})((T - W)/((1 - O) * W) + 1)$  columns, where  $T$  is the length of the signals (in time units).

We applied Gaussian kernel with width parameter 0.25, overlap 0.0-0.3 and used sliding windows between 200 and 500 days. Fig. 2 depicts how the CC time lags develop over monitoring time. NGC 5548 both lines shows continuous symmetrical change of CC time lags over time, with larger amplitude in the case of the H $\beta$  line. The largest and smallest variation for both lines occurred within time window corresponding to the highest state and lowest state of activity (compare Fig.1 and

Fig.2). However, we find that MI method applied on both lines of NGC 5548 (Fig.7), as well as for other cases, fails, as it only infers the results for time windows with enough points (dark color in the Fig.7).

Contrary to the NGC 5548, the both lines of Arp 102B and H $\beta$  lines of 3C 390.3 and E1821+643 exhibit abrupt 'regime shift': from nonsignificant lags (several days) up to about magnitude of hundred days. The magnitude of shift is larger in the case of H $\beta$  than H $\alpha$  line of Arp 102B, while in the case of 3C 390.3 and E1821+643 we did not have enough data for H $\alpha$  line and H $\gamma$  analysis, since it has been required by method at least of 50 points in each sliding window.

It is clearly seen that abrupt change in time evolution of 3C 390.3 lags (Fig.2) corresponds to the highest state of both light curves (Fig. 1), while in the case of E1821+643, the regime shift corresponds to the highest state of continuum only. In the case of Arp 102 B 'the regime shift' of time lag evolution corresponds to the prominent feature of the H $\beta$  line around MJD 54000 (see Fig.3 in Shapovalova et al. (2013)).

In the case of NGC 4051, the almost diagonal patches are occupied displaying tendency that time lag evolve linearly over time. In the case of other time series, the absence of such pattern could display tendency for non-linear time delay evolution. Both light curves of NGC 4051 (see Fig. 3 in Denney et al. (2009)) display almost oscillatory behavior without prominent changes.

To obtain a quantitative understanding of the feasibility of above mentioned results, we modeled a five artificial cases using synthetic light curves obtained by the Cholesky decomposition technique implemented in JAVELIN code (Zu et al. 2011). One virtue of JAVELIN is that it produces an explicit mean model light curves constrained by the data.

Since the structures shown in the Fig. 2 (c, d panels) exhibits strange **behavior**, the lags change from something close to zero to a hundred days, we constructed 10 to 30 times denser synthetic light curves than original ones by taking into account

the length of monitoring baselines corresponding to original length of monitoring campaigns and cadences of daily or lower value.

For each case we generated 50 MCMC realizations of light curves assuming that fictive objects posses variability parameters of our 5 objects: variance  $\sigma[mag]$  and DRW time scales of  $\tau_D[days]$ . The panels in the Fig. 3 and Fig. 4 show the Case 1, 2, 3 and Case 4, 5 mean model light curves respectively while the results of application of Gaussian kernel based method on those synthetic light curves can be seen in the Figs. 5 and 6. Time evolved Gaussian kernel lags (c, d panels in the Fig. 5) of synthetic light curves (see the middle panel in the Fig. 3) cloned from the mother curves of Arp 102B, do not exhibit such changes as those seen in the Fig. 2 (c and d panels). **Clearly, in the case of Arp 102B clones, temporal non variability of time lags is a consequence of stationarity and constant lags of those curves. However, other cases of cloned light curves show strong variation in lags. In the case cases of cloned light curves of NGC 5548, 3C390.3, E1821+643 we notice that DRW time scales are larger than used sliding time windows of the Gaussian method, while in the case of NGC 4051 those values are close to each other ( $\tau_D = 3.8$  and  $W \sim 10$ ). These anomalies can rise the instability of Gaussian kernel based method which induces significant variations of time lags of those cloned curves. From the other hand it seems that DRW time scales of Arp 102B light curves, have 'Goldilocks' values, and the method shows more stability.**

Moreover, all synthetic cases exhibit the extremum values of time lags within the time range corresponding to the shocks in time series.

The static time lags derived by standard slotting and model technique (ZDCF and SPEAR, Kovačević et al. (2014)) are given in Table 3. The magnitude of lag time evolution for all our cases over monitoring periods are within  $3\sigma$  interval of values obtained by standard techniques (Table 3).

Since many authors reported individual time lag determinations of our objects (compiled in Kovačević et al. (2014)) with the exception of E1823+643, we were

able to calculate their weighted mean in order to have a more informative comparison to the evolved time lags. As the time lags are accompanied with asymmetric errors, their weighted mean and error combinations can be found by Model 1 reported in Barlow (2003). This Model provides convergence of iterative numerical solution. If values and their asymmetric errors are given as  $(x_i, \sigma_i^-, \sigma_i^+) \in f_i$ , where  $f_i$  is a distribution, the combination of errors is done by finding the variance  $V_i = 0.25(\sigma_i^+ + \sigma_i^-)^2 + 0.25(\sigma_i^+ - \sigma_i^-)^2(1 - \frac{2}{\pi})$  and skew  $\gamma_i = \frac{1}{\sqrt{2\pi}}[2(\sigma_i^{+3} - \sigma_i^{-3}) - \frac{3}{2}(\sigma_i^+ - \sigma_i^-)(\sigma_i^{+2} - \sigma_i^{-2}) + \frac{1}{\pi}(\sigma_i^+ - \sigma_i^-)^3]$  for each contributing distribution separately. Adding these up gives the cumulants  $V$  (variance) and  $\gamma$  (skew) of the combined distribution function. If we write  $D = \sigma^+ - \sigma^-$  and  $S = \sigma^{+2} + \sigma^{-2}$  then the combined errors  $\sigma^+, \sigma^-$  are obtained numerically from following nonlinear equations

$$\begin{aligned} S &= 2V + \frac{D^2}{\pi} \\ D &= \frac{2}{3S}(\sqrt{2\pi}\gamma - D^3(\frac{1}{\pi} - 1)) \end{aligned} \quad (8)$$

From the equations above one can find the solution for  $D$  and  $S$ , after that simple algebraic manipulation leads to the  $\sigma^+, \sigma^-$ . As for an unbiased weighted mean estimator one has to take following expression

$$\check{x} = \frac{\sum_{i=1}^n \omega_i(x_i - bi)}{\sum_{i=1}^n \omega_i} \quad (9)$$

where  $b_i = \frac{\sigma_i^+ - \sigma_i^-}{\sqrt{(2\pi)}}$  and weights are given as  $\omega_i = \frac{1}{V_i}$ . We calculated the weighted mean of time lags obtained by other authors ( $\check{\tau}$ ) and combined their errors in the form of  $\sigma^+$  and  $\sigma^-$  according to Barlow method (see Table 4). Their values agree with those time evolved. Namely, the time evolved lags  $\tau$  are within the region  $[\sigma^+ + \check{\tau}, \sigma^- - \check{\tau}]$  of weighted mean lags. There is only slight difference in the case of time lags of Arp 102B.

| Object    | LC         | $TS_{nc}$ | nc (1%) | nc(5%) | nc(10%) |
|-----------|------------|-----------|---------|--------|---------|
| E1821+643 | cnt        | -0.17     | -2.6    | -1.94  | -1.6    |
|           | H $\beta$  | -0.15     | -1.94   | -1.94  | -1.61   |
| Arp 102B  | cnt        | -0.42     | -2.6    | -1.9   | -1.6    |
|           | H $\alpha$ | -0.95     | -2.6    | -1.9   | -1.6    |
| NGC4051   | H $\beta$  | -0.72     | -2.6    | -1.9   | -1.6    |
|           | cnt        | -0.6      | -2.6    | -1.9   | -1.6    |
| 3C 390.3  | H $\beta$  | -0.4      | -2.6    | -1.9   | -1.6    |
|           | cnt        | 0.8       | -2.6    | -1.9   | -1.6    |
| NGC 5548  | H $\beta$  | 0.3       | -2.6    | -1.9   | -1.6    |
|           | cnt        | -0.650    | -2.6    | -1.9   | -1.6    |
|           | H $\alpha$ | -1.3      | -2.6    | -1.9   | -1.6    |
|           | H $\beta$  | -1.5      | -2.6    | -1.9   | -1.6    |

Table 2: The Augmented Dickey-Fuller statistics for the test of non stationarity. In the column LC are given light curves;  $TS_{nc}$  column provides test statistics involving no constant trend ; the column nc gives critical values for the test statistic at the 1%, 5%, 10% levels. In all cases the null-hypothesis that process is non-stationary is satisfied.

### 3 Conclusion

To conclude, the present work is a continuation of our previous publication (Kovačević et al. (2014)) on time series analysis of enlarged data set: Arp 102B, 3C 390.3, NGC 5548, NGC 4051 and in addition E1821+643. Here we examined both non-stationarity of their time series by means of ADF and time evolution of time lags between continuum and line curves by means of Gaussian kernel based CC and MI. We found strong evidence for non-stationarity of all time series in ADF test regime without constant terms, which corresponds to modeling damped random walk.

| Object    | Line       | $\tau_{SPEAR}$          | $\tau_{ZDCF}$                                   |
|-----------|------------|-------------------------|---|
| Arp 102B  | H $\alpha$ | 23.8 $^{+27.5}_{-18.8}$ | 15.0 $^{+24.4}_{-13.8}$                         |
|           |            |                         | 157.1 $^{+20.3}_{-110.1}$                       |
| 3C 390.3  | H $\beta$  | 47.6 $^{+57.2}_{-37.2}$ | 22.8 $^{+64.0}_{-20.9}$                         |
|           |            |                         | 76.9 $^{+79.1}_{-74.7}$ 94.5 $^{+27.1}_{-48.0}$ |
| NGC 5548  | H $\alpha$ | 43.5 $^{+47.7}_{-39.3}$ | 27.0 $^{+14.4}_{-5.7}$                          |
|           | H $\beta$  | 45.4 $^{+47.0}_{-43.0}$ | 49.2 $^{+18.6}_{-7.7}$                          |
| NGC 4051  | H $\beta$  | 2.8 $^{+3.1}_{-2.3}$    | 2.6 $^{+0.9}_{-1.1}$                            |
| E1821+643 | H $\beta$  | 125.6 $^{+3.4}_{-2.3}$  | 26.3 $^{+48.2}_{-24.1}$                         |
|           |            |                         | 137 $^{+54}_{-19}$                              |

Table 3: The time lags determined by ZDCF and SPEAR method (see Kovačević et al. (2014)).

| Object   | $\check{\tau}$ | $\sigma^+$ | $\sigma^-$ |
|----------|----------------|------------|------------|
| Arp 102B | 14.4           | 39.1       | 27.4       |
| 3C 390.3 | 51.4           | 77.1       | 76.9       |
| NGC 5548 | 10.8           | 47.5       | 48         |
| NGC 4051 | 2.2            | 8.1        | 9.1        |

Table 4: The weighted mean of time lags ( $\check{\tau}$ ) obtained by other authors (compiled in Kovačević et al. (2014)) and their combined errors ( $\sigma^+$ ,  $\sigma^-$ ), calculated by method of Barlow (2003).

The application of Gaussian kernel based CC through observed data set shows specific temporal behavior of time lags between continuum and lines. Running this method to surrogate light curves (which are stationary and having constant lags) the results still show strong variation in lags (excluding the case of cloned curves of Arp 102B). In those surrogate cases with significantly variable time lags, the certain discrepancies between the DRW time scales and used sliding time windows of Gaussian kernel based CC are noticed. So the recorded time lag variations in the

cases of NGC 5548, 3C 390.3, NGC 4051 and E1821+643 can be a consequence of instability of the method due to above mentioned anomalies. From the other hand, it seems that DRW time scale of Arp 102B has 'Goldilocks' value so the method shows more stability (time lags of its cloned curves do not vary with time notably) and detected variations could correspond to the real behavior of time lag over observed time period. Nevertheless, tracking of MI time evolution of time lags has failed due to a small number of points in used time windows. The results highlight specific issues so the further analysis is necessary.

**Acknowledgements** This work was supported by the Ministry of Education , Science and Technological Development of Republic of Serbia through the project Astrophysical Spectroscopy of Extragalactic Objects (176001) and RFBR (grants N12-02-01237a, 12-02-00857a) (Russia) and CONACYT research grants 39560, 54480, and 151494 (Mexico)..

## References

- [1] Andrae, R., Kim, D.W., and Bailer-Jones, C. A. L., Assessment of stochastic and deterministic models of 6304 quasar light curves from SDSS Stripe 82, *A & A*, 554, 137, (2013)
- [2] Babu, P., Stoica, P., Bos, R., Spectral analysis of nonuniformly sampled data - a review, *Digit. Signal Process.*, 20, 359378, doi:10.1016/j.dsp.2009.06.019, (2009)
- [3] Barlow, R., Asymmetric Systematic Errors, arXiv:physics/03061381v1, (2003).
- [4] Bjoernstad, O. N. and Falck, W., Nonparametric spatial covariance functions: Estimation and testing, *Environ. Ecol. Stat.*, 8, 5370, (2001)
- [5] Broersen, P. M., de Waele, S., The Accuracy of Time Series Analysis for Laser-Doppler Velocimetry, in: Proceedings of the 10th International Symposium on Applications of Laser Techniques to Fluid Dynamics, 10-13 July Lisbon, Portugal, 1-10, (2000)

- [6] Brooks, C., *Introductory Econometrics for Finance* 3rd Edition, Cambridge University Press, Cambridge, (2014)
- [7] Chatfield, C., *The analysis of time series: an introduction*, Series:Texts in statistical science, CRC Press, Florida, US, 6th Edn., (2004)
- [8] Cackett, E. M., Horne, K.: 2005, Photoionized H $\beta$  emission in NGC 5548: It breathes!, *MNRAS*, 365, 4, 1180-1190, (2005)
- [9] Czerny, B., Hryniwicz, K., Eddington ratio, in *Fifty Years of Quasars: From Early Observations and Ideas to Future Research*, eds D' Onforio. M, Marziani, P., Sulentic, J., Springer, 319-323, (2012)
- [10] Cover, T. M. , Thomas, J. A., *Elements of Information Theory*, Wiley, New York, (1991)
- [11] Denney, K. D., Bentz, M. C., Peterson, B. M., Pogge, R. W., et al., The Mass of the Black Hole in the Seyfert 1 Galaxy NGC 4593 from Reverberation Mapping ,*ApJ*, 653, Issue 1, 152-158, (2006)
- [12] Denney, K. D., Watson, L. C., Peterson, B. M., Pogge, R. W., Atlee,et al, A Revised Broad-line Region Radius and Black Hole Mass for the Narrow-line Seyfert 1 NGC 4051, *ApJ*, 702, Issue 2, 1353-1366, (2009)
- [13] Dietrich, M., Kollatschny, W., Optical and ultraviolet emission-line variability of NGC 5548 - The coordinated UV and optical monitoring campaign of 1989 *A&A*, 303, 405, (1995)
- [14] Edelson, R. A., Krolik, J. H.,The discrete correlation function - A new method for analyzing unevenly sampled variability data, *ApJ*, 333, 646-659, (1988)
- [15] Gaskell, C. M., Sparke, L. S., Line variations in quasars and Seyfert galaxies, *ApJ*, 305, 175-186, (1986)
- [16] Goad M. R., Korista, K. T, Knigge, C., Is the slope of the intrinsic Baldwin effect constant?, *MNRAS*, 352, 277-284, (2004)
- [17] Gilbert K. M., Peterson B. M., An intrinsic Baldwin effect in the H $\beta$  broad emission line in the spectrum of NGC 5548 2003, *ApJ*, 587, 123-127, (2003)

- [18] Kelly, B. C., Bechtold, J., & Siemiginowska, A., Are the variations in quasar optical flux driven by thermal fluctuations?, *ApJ*, 698, 895-910, (2009)
- [19] Kinney, A. L., Rivolo, A. R., Koratkar, A. P., A study of the Baldwin effect in the IUE data set, *ApJ*, 357, 338-345, (1990)
- [20] Kolmogorov, A. N., Logical basis for information theory and probability theory. *IEEE Trans. Information Theor.*, 14, 662664, (1968)
- [21] Korista K. T., Goad M. R., What the optical recombination lines can tell us about the broad-line regions of active galactic nuclei, *ApJ*, 606, 749-762, (2004)
- [22] Koratkar, A. P., Gaskell, C. M.: Structure and kinematics of the broad-line regions in active galaxies from IUE variability data *ApJS*, 75, 719-750, (1991)
- [23] Kovačević, A., Popović, L. Č., Shapovalova, A. I., Ilić, D., Burenkov, A. N., Chavushyan, V., Time series analysis of active galactic nuclei: The case of Arp 102B, 3C 390.3, NGC 5548 and NGC 4051, *ASR*, 54, 1414-1428, (2014)
- [24] Kozłowski, S., Kochanek, C. S., Udalski, A., Wyrzykowski, L., Soszynski, I., et al., Quantifying quasar variability as part of a general approach to classifying continuously varying sources, *ApJ*, 708, 927-945, (2010)
- [25] Krolik, J. H., Horne, K., Kallman, T. R., Malkan, M. A., Edelson, R. A., Kriss, G. A., UV Variability of NGC 5548: Dynamics of the Continuum Production Region and Geometry of the Broad Line Region, *ApJ*, 371, 541-562, (1991)
- [26] Lomb, N. R., Least-Squares Frequency Analysis of Unevenly Spaced Data, *Astrophysical and Space Science*, 39, 447-462, (1976)
- [27] MacLeod, C. L., Ivezić, Ž., Kochanek, C. S., Kozłowski, S., Kelly, B., et al., Modeling the time variability of SDSS stripe 82 quasars as a damped random walk, *ApJ*, 721, 1014-1033, (2010)
- [28] Mayo, W.: Spectrum measurements with laser velocimeters (in: Proceedings of the Dynamic Flow Conference 1978), in: Selected Papers on Laser Doppler Velocimetry, ed by: Adrian, R. J., Society of Photo-Optical Instrumentation Engineers: SPIE milestone series; 78, chap. 3, 222235, SPIE Press, (1993)

- [29] Neyman, J., Pearson, E.S., On the problem of the most efficient tests of statistical hypotheses. *Philos. Trans. R. Soc. A*, 231,289337, (1933)
- [30] Pessah, M. E., Mass measurements of AGN from multi-Lorentzian models of X-ray variability. I. Sampling effects in theoretical models of the  $\sigma_{rms}^2$ - $M_{BH}$  correlation, *ApJ*, 655, 66-76, (2005)
- [31] Pogge, R. W., Peterson, B. M., The intrinsic nature of the Baldwin effect, *AJ*, 103, 1084-1088, (1992)
- [32] Rehfeld, K., Marwan, N., Heitzig, J., Kurths, J., Comparison of correlation analysis techniques for irregularly sampled time series, *Nonlin. Processes Geophys.*, 18, 389404, doi:10.5194/npg-18-389-2011, (2011)
- [33] Rehfeld, K., Kurths, J., Similarity estimators for irregular and age-uncertain time series, *Clim. Past*, 10, 107-122, (2014)
- [34] Said, S. E., Dickey, D., Testing for Unit Roots in Autoregressive Moving-Average Models with Unknown Order, *Biometrika*, 71, 599-607, (1984)
- [35] Scargle, J., Studies in astronomical time series analysis, I. Modeling random processes in the time domain, *ApJS.*, 45, 171, (1981)
- [36] Scargle, J., Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data., *ApJ*, 263, 835853, (1982)
- [37] Shannon, C. E., Weaver, W., *The Mathematical Theory of Communication*, University of Illinois, Urbana, IL, (1949)
- [38] Shapovalova, A. I., Doroshenko, V. T., Bochkarev, N. G., Burenkov, A. N., Carrasco, L., Chavushyan, V. H. et al., Profile variability of the H $\alpha$  and H $\beta$  broad emission lines in NGC 5548, *A&A*, 422, 925-940, (2004)
- [39] Shapovalova, A. I., Popović, L. Č., Burenkov, A. N., Chavushyan, V. H., Ilić, D., et al., Spectral optical monitoring of 3C 390.3 in 1995-2007. I. Light curves and flux variation in the continuum and broad lines , *A&A*, 517, A42, 1-27, (2010)

- [40] Shapovalova, A. I., Popović, L. Č., Ilić, D., Burenkov, A. N., Chavushyan, V. H., Ilić, D., Kollatschny, W. et al, Spectral optical monitoring of a double-peaked emission line AGN Arp 102B: I. Variability of spectral lines and continuum, *A&A*, 559, A10, 1-17, (2013)
- [41] Shapovalova, A. I., Popović, L. Č., Burenkov, A. N., Chavushyan, V. H. et al., manuscript in preparation, (2015)
- [42] Shawe -Taylor, J., Cristianini, N., Kernel methods pattern analysis, Cambridge University Press, (2004)
- [43] Stoica, P., Li, J., He, H., Spectral analysis of nonuniformly sampled data: A new approach versus the periodogram, *IEEE T. Signal Process.*, 57, 843858, doi:10.1109/TSP.2008.2008973, (2008)
- [44] Zu, Y., Kochanek, C., S., & Peterson, B. M., An alternative approach to measuring reverberation lags in Active Galactic Nuclei, *ApJ*, 735, Issue 2, article id. 80, 1-16, (2011)
- [45] Zu, Y., Kochanek, C. S., Kozlowski, S. & Udalski, A., Is quasar optical variability a damped random walk?, *ApJ*, 765, id106, 1-7, (2013)

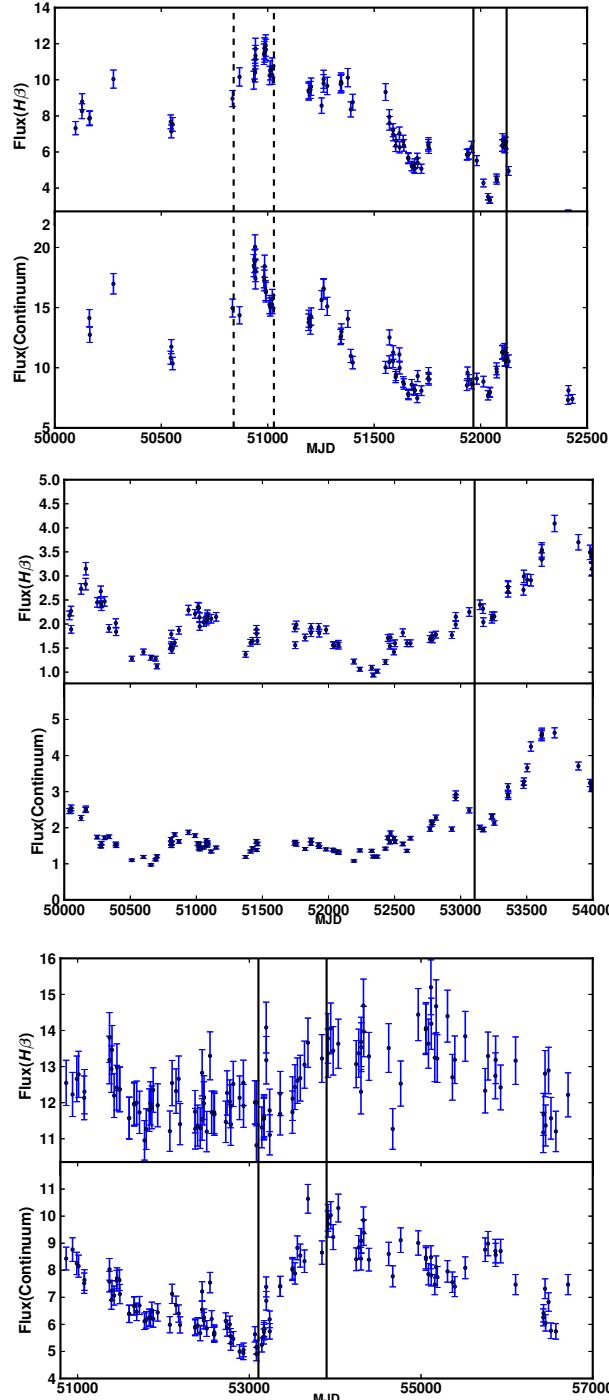


Fig. 1: Upper plot: Time series from our spectral monitoring program data on NGC 5548 for optical continuum flux ( $10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$ , at 5190 Å) and H $\beta$  emission line flux ( $10^{-13} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$ ) in 1996-2002. Vertical dashed and solid lines mark the highest (1998 Jun 26; JD 2450991) and lowest (2002 Jun 4; JD 2452430) activity period of the light curves respectively. Middle and bottom plot: The same as in upper plot but for spectral monitoring of 3C 390.3 in 1995-2007 and E1821+643 in 1998-2014. Solid lines mark the highest activity period of the light curves.

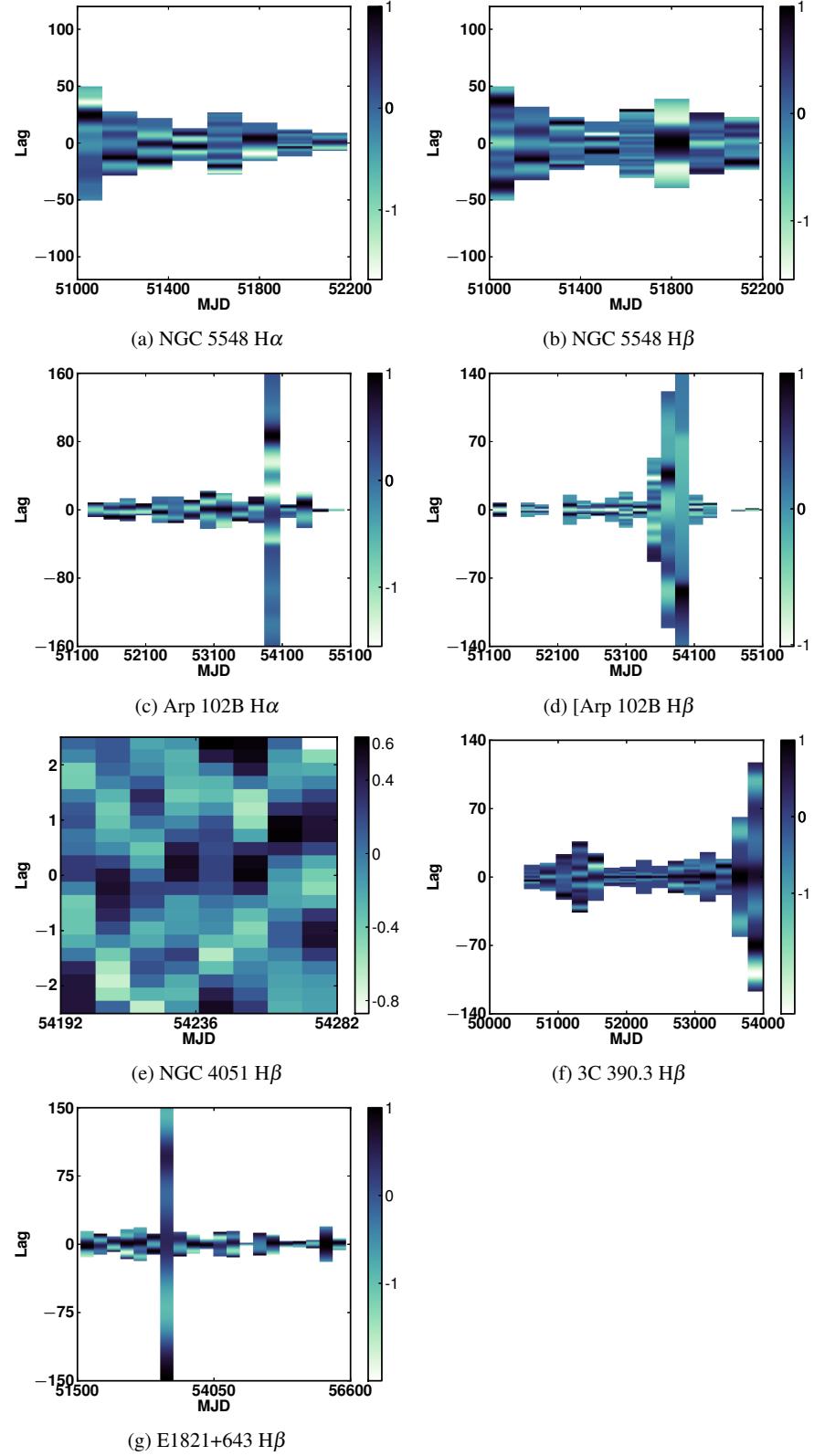


Fig. 2: Time evolution of the CC time lags for the time periods of monitoring campaigns. Note that NGC 5548 exhibits clearly symmetrical pattern of behavior, while the largest values of CC time lags of NGC 4051 develops over time almost diagonally. No symmetry or diagonally distributed elements is seen in the case of Arp 102B, 3C 390.3 and E1821+643. Colorbars represent the values of Gaussian kernel based CC coefficients.

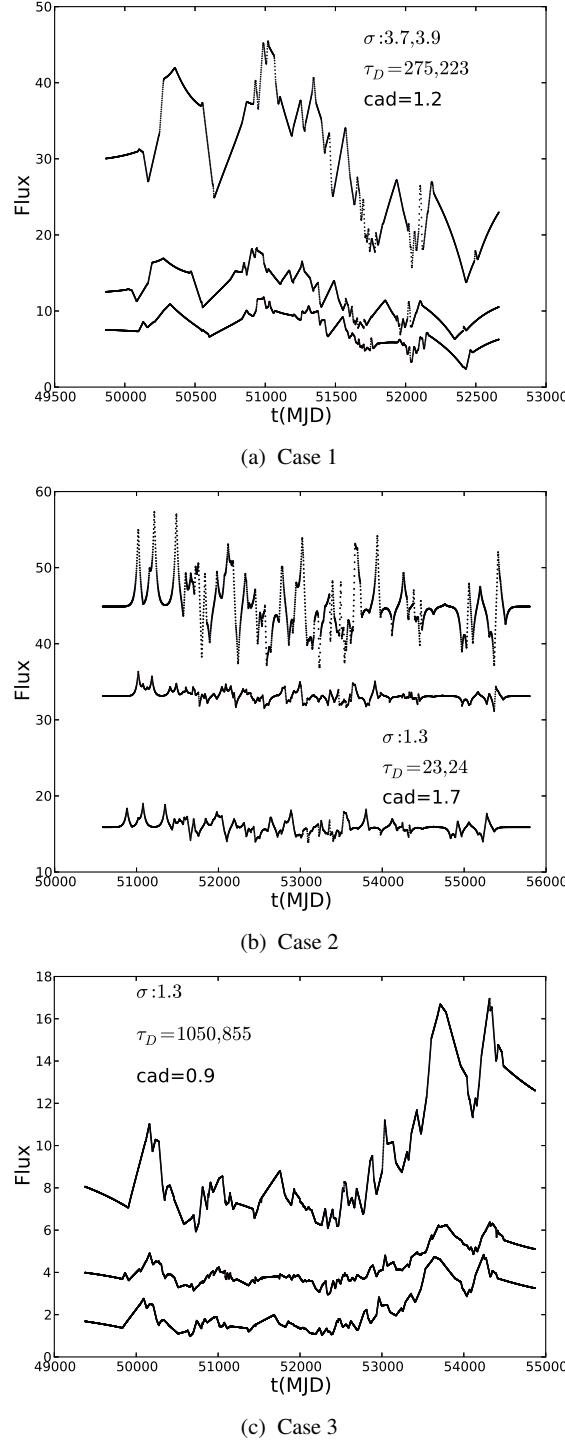
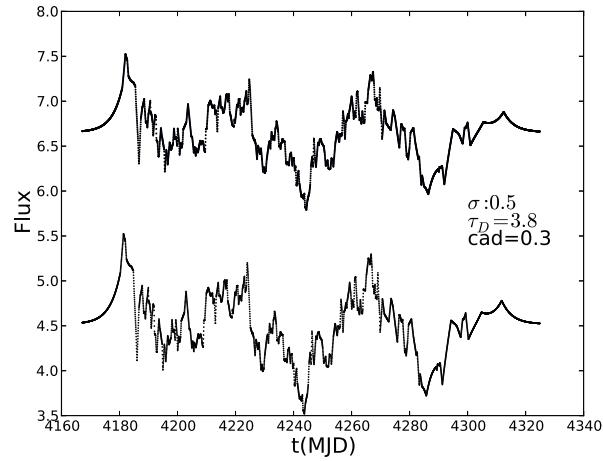
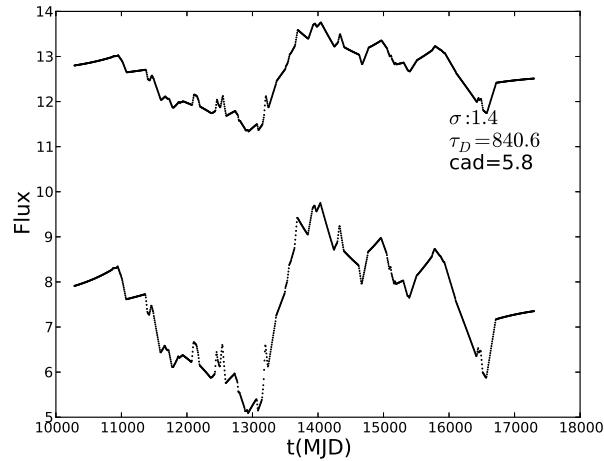


Fig. 3: Synthetic light curves generated using the variability parameters  $\sigma$ ,  $\tau_D$  and cadence, listed on each panel, in units of mag, days, and days respectively. Assumed objects possessed the variability parameters of NGC 5548 (Case 1, upper plot), Arp102B (Case 2, middle plot) and 3C 390.3 (Case 3, bottom plot). All synthetic light curves cover original length of monitoring campaigns. On each panel the upper and middle curve mimics emission lines while the bottom curve mimics continuum.



(a) Case 4



(b) Case 5

Fig. 4: The same as in Fig. 4 but for the NGC 4051 (Case 4, top plot) and E1821+643 (Case 5, bottom plot).

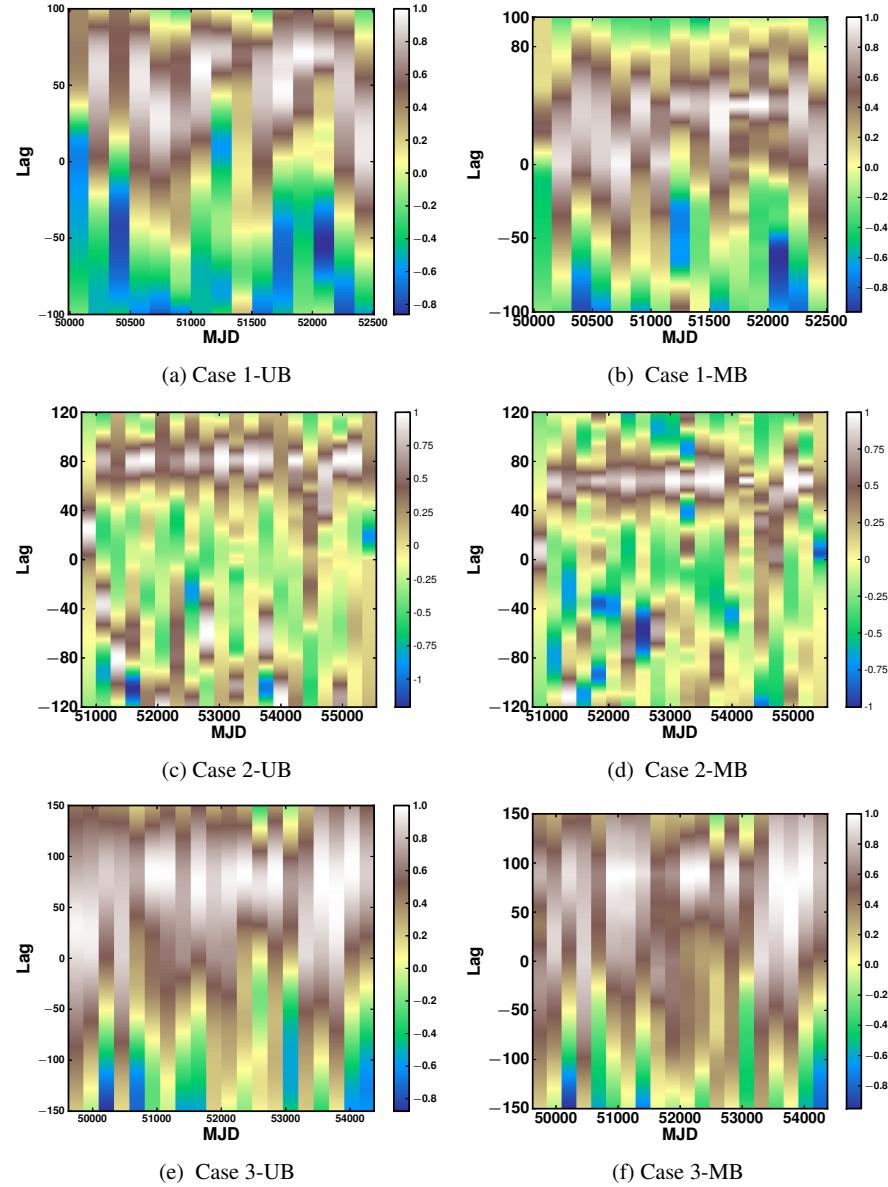


Fig. 5: Gaussian kernel based method applied on synthetic light curves Cases 1, 2, 3 (see Fig. 3). Left panels: UB stands for the time evolution of lags between the upper and bottom light curve from Fig.3. Right panels: MB stands for time evolution of lags between middle and bottom light curves from Fig. 3.

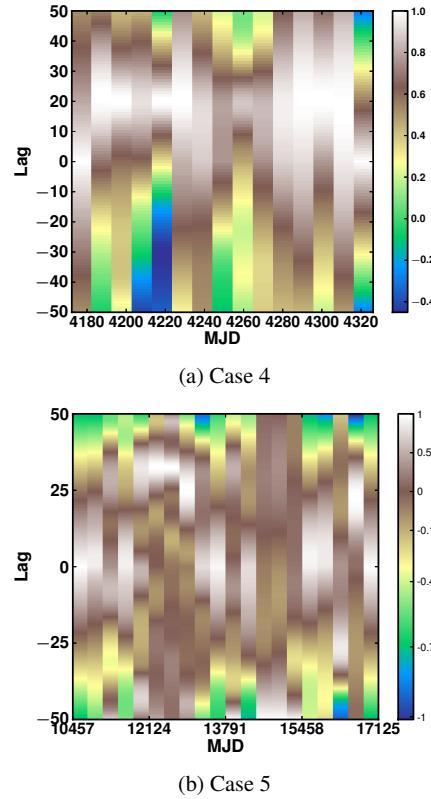


Fig. 6: Same as in Fig. 5 but for Cases 4 and 5 (see synthetic light curves given in the Fig. 4).

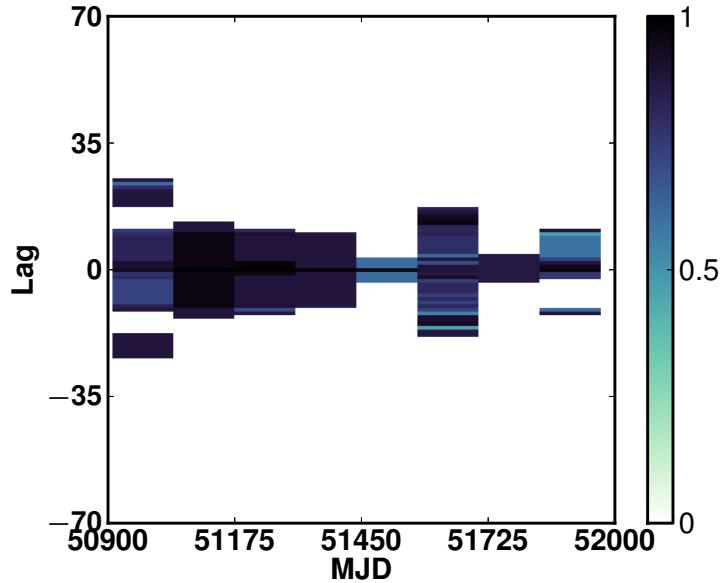


Fig. 7: The time evolution of MI based time lag between continuum and H $\beta$  line of NGC 5548.