MNRAS 000, 1–4 (2015) Preprint 12 February 2016 Compiled using MNRAS LaTeX style file v3.0

Solving the puzzle of discrepant variability on monthly time scales implied by SDSS and CRTS datasets

Krzysztof Suberlak,[1⋆] Željko Ivezić, [1] Chelsea L. MacLeod,[2] Matthew Graham,[3] Branimir Sesar[4]
[1]Department of Astronomy, University of Washington, Seattle, WA, United States
[2]Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh, United Kingdom
[3]Center for Data-Driven Discovery, California Institute of Technology, Pasadena, CA, United States
[4]National Optical Astronomy Observatory, Tucson, AZ, United States.

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We present improved error analysis for the 3,800 CRTS (Catalina Real-Time Transient Survey) optical quasar light curves from the Sloan Digital Sky Survey Stripe 82 catalog. SDSS imaging survey has provided a time-resolved photometric dataset which greatly improved our understanding of the quasar optical continuum variability: data for monthly and longer timescales are consistent with a damped random walk. Recently, newer data obtained by CRTS (Catalina Real-Time Transient Survey) provided puzzling evidence for enhanced variability, compared to SDSS results, on monthly time scales. Quantitatively, SDSS results predict about 0.06 mag rms variability for timescales below 50 days, while CRTS data show about a factor of two larger root-mean-square for spectroscopically confirmed SDSS quasars. Our analysis presented here has successfully resolved this discrepancy as due to slightly underestimated photometric error estimates provided by the CRTS image processing pipelines. The photometric error correction factors, derived from detailed analysis of non-variable SDSS standard stars that were re-observed by CRTS, are about $20-30\%$, and result in a quasar variability behavior implied by the CRTS data fully consistent with earlier SDSS results.

1 INTRODUCTION

Quasar variability is an important characteristic that has a potential to shed light on the structure of the innermost region of the accretion disk, and has been the subject of research for the past half century (Matthew, Sandage 1963, Vanden Berk + 2004, Grier+2012, Kozlowski+2016). The Sloan Digital Sky Survey (SDSS, Schmidt+2010, Sesar+2007) and Catalina Real-Time Transient Survey (CRTS, Djorgovski+2012, Drake+2009) allowed an unprecedented study of well-calibrated light curves.

Assuming that the observed variability of a lightcurve - discretely sampled time series - is caused by a continuous underlying physical phenomenon, to characterise it we can use a mathematical formalism of a stochastic process model (Kasliwal+2015). A one-parameter Auto-Regressive, AR(1), process may correspond to the physical situation of an environment where a disturbance is diffused and returns to the median value. Various explanations exist for the nature of such disturbances, including supernovae, microlensing, accretion disk instabilities, and thermal fluctuations (Kelly+2009, Kelly+2011, Collier&Peterson 2001, Ruan+2014). Most likely, thanks to Reverberation Mapping studies (Peterson 2004, Fausnaugh+2016, Munoz+2016), we understand the accretion disk to be the main source of the underlying variability.

At the cadence of the CRTS survey (more than few days), the AR(1) process (Damped Random Walk model) describes well the quasar variability (Butler Bloom 2011, MacLeod+2010,2011,2012, Kelly+2009, Zu+2011,2013, Kozlowski+2010a, Ruan+2012). [A deviation from the DRW model at very short timescales, reported by Kasliwal+2015, is based on Kepler lightcurves at cadences shorter than a day, which is below the observational limit of CRTS. For this reason, in this paper DRW is assumed to be a good descriptor of the quasar stochastic variability].

The DRW model has two parameters - an amplitude of variability, and a characteristic timescale, that corresponds to the timescale of damping of thermal oscillations or orbital timescale (Kelly+2009). When the lightcurves are sparsely sampled, it is more meaningful to describe their ensemble properties, using the structure function (SF), which is well defined for a collection of variable objects driven by the DRW (Kozlowski+2010, 2016, Vanden Berk+2004, Schmidt+2010, Hawkins+2002)

Recent SDSS-based studies (MacLeod+2010, Kelly+2009) reflect the traditional timescale of $\tau > 100$ days in quasar rest frame, supported by OGLE results of

Zu+2014 with $17 \leq \tau \leq 2700$ days. However, a CRTS-based study of Graham+2014 that used Slepian Wavelet Variance methodology found τ 54 days. The latter is consistent with the Kepler-based results of Kasliwal+2015 for cadences above a day, who found $17 \leq \tau \leq 2700$ days. The aim of this paper is to revisit the analysis of the CRTS quasars using a well-tested approach of the Structure Function, and detailed error analysis (MacLeod+2010, MacLeod+2012, Simonetti+1984, Vanden Berk+2004, Sumi+2005, Bauer+2009).

In section 2 we describe the properties of our dataset, all selection criteria, and the structure function analysis. In section 4 we explain the results of our analysis of the small timescale subset of the structure function. This is followed by discussion of the impact of our result, and conclusions in sections **??** and 5.

2 DATA ANALYSIS

Catalina Real-time Transient Survey (CRTS) employed three telescopes ("0.7 m Catalina Sky Survey Schmidt and 1.5 m Mount Lemmon Survey telescopes in Arizona, and the 0.5 m Siding Springs Survey Schmidt in Australia" Graham+2015), each with a 4kx4k CCD detector (Djorgovski+2011). All observations were taken in an unfiltered white light, and calibrated to the Johnson system V-band zero point (Drake+2013). In this research we use CRTS lightcurves of 7932 spectroscopically-confirmed QSOs and 52133 standard stars from the equatorial region of the sky ($22h24m$ < R.A. < $04h08m$ and |Dec| < $1.27deg$) known as the Stripe 82 (Sesar+2010). Our sampling criterion : more than 10 measurements per lightcurve, leaves 7707 quasars and 49385 stars. To improve signal to noise we perform day-averaging, and thus we also require more than 10 individual days of observations per lightcurve, leaving in the final sample 7601 quasars and 48250 stars (see Fig.2 for statistical description of the quasar sample).

We complement the CRTS data for each object with the SDSS photometry by creating a positionally cross-matched catalog using astropy `match_to_catalog_sky` routine. For catalog matching we used the SDSS Stripe 82 standard stars catalog ver. 2.6 (Ivezic+2007), and the SDSS DR7 Stripe 82 quasars catalog (Abazajian+2009). Since the SDSS star catalog contains accurate data for 1006849 stars , all CRTS stars were found to have a matching SDSS counterpart within 0.01 arcsec. Of 7601 CRTS quasars, 7586 had an SDSS counterpart within 0.36 arcsec. The remaining 15 CRTS quasars had an SDSS counterpart within between 2 and 9 arcmin [- they were ignored ?]

To improve the signal-to-noise ratio we performed day-averaging of light curves. The day-magnitude is the mean of individual measurements, and the day-error e_w is the weighted sum of individual errors e_i:

$$e_w = \frac{1}{\sqrt{\sum_i w_i}} \tag{1}$$

where $w_i = 1/e_i^2$ is the weight assigned to each point. To avoid unphysically small error values, we added 0.01 *mag* to e_w in quadrature if $e_w < 0.02$ *mag*. [A typical sampling interval of day-averaged CRTS Quasar lightcurves is ≈ 20 days.]

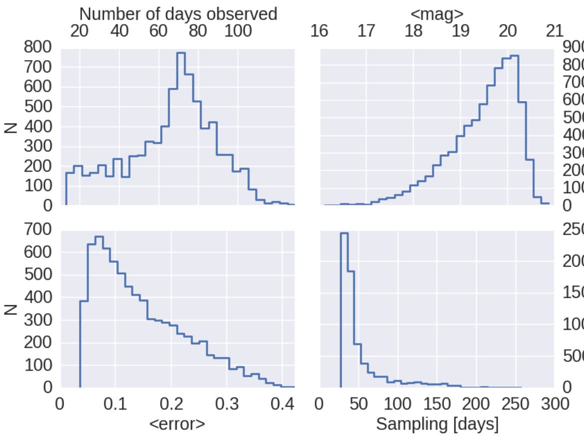

Figure 1. Statistical information about the CRTS QSO final sample of 7601 day-averaged light curves. The upper-left panel shows the number of days that a given quasar was observed, i.e. $max(MJD) - min(MJD)$ per light curve. The upper-right panel shows the average light curve magnitude. The bottom-left panel shows the light curve averaged error. If i, $i+1$ are two consecutive day-averaged observations in the light curve, then $MJD_i - MJD_{i+1}$ is the sampling interval. The bottom-right panel shows light curve - averaged sampling intervals. The light curve and sample-averaged mean and median error is 0.22. The mean brightness is 19.50, and the mean number of individual observations per light curve is 209. Within that sample , 96% observations of quasars span the time of 7-9 years. 91.2% of quasars were observed between 1 to 4 times per night.

3 STRUCTURE FUNCTION

The structure function (SF) is a well-studied approach to characterising quasar light curves. It describes the relationship between the time lag and the amplitude of brightness variability (Cristiani+1996, Schmidt+2010, Vanden Berk +2004, de Vries + 2005, Rengstorf + 2006).

Similarly to Schmidt+2010, we calculate the SF for all objects using the magnitude difference $\Delta m_{i,j}$ between light curve points i and j, separated by a time lag $\Delta t_{i,j}$. To avoid the additional uncertainty of redshift estimates based on the SDSS spectra, we also use time lags in the observed frame. We add the error information in quadrature: $e_{ij} = \sqrt{e_i^2 + e_j^2}$. Such 'master file' containing $\Delta m_{i,j}$, $\Delta t_{i,j}$ and e_{ij} was made for each object.

To analyze the ensemble $\Delta m_{i,j}$ variability we bin the $\Delta m_{i,j}$ points from all objects in a given sample in logarithmic $\Delta t_{i,j}$ space. The number of chosen bins is a choice of convenience between very coarse grid (a small number of bins) or a very fine grid, risking a small number of points per bin. We find that from a choice of N =50, 100, 200, and 400 bins , $N = 200$ is the right choice between computational efficiency and the information content.

We calculate the SF using an approximate prescription involving marginalizing the log-likelihood of the probability in $p(SF)$ and $p(\mu)$ space per bin, after [Ivezic+2013]

$$\frac{dp(SF}{dSF}\bigg|_{SF=mode} = 0 \tag{2}$$

For each $\Delta t_{i,j}$ bin we calculate four statistical descriptors, shown on Fig. 3 : standard deviation from the mean (σ, Gaussian robust deviation from the mean $\sigma_G = 0.7414(q_{75} - q_{25})$, the SF and the mean ($\mu$).

In plotting the CRTS quasar SF we add the fiducial DRW model to allow a quick comparison between the DRW model and the data:

$$SF(\Delta t) = SF_\infty \cdot \left(1 - e^{-\Delta t/\tau}\right)^{1/2} \tag{3}$$

with the model error $\sqrt{SF(\Delta t)^2 + err_{SF}^2}$.

4 RESULTS

To study the intrinsic variability of quasars we use the standard stars that do not exhibit any intrinsic variability as the control sample. We define our sample in a following way: for both stars and quasars, we require SDSS r-band $17 < m_{SDSS,r} < 19$, and the CRTS lightcurve-averaged error to be $\langle e_w \rangle < 0.3$. We further subdivide this sample into three magnitude bins: 17-18, 18-18.5, and 18.5-19 mag. To acknowledge the influence of colors on photometric error, we divide the stars in two color bins: "red" with $1 < g - i < 3$ and "blue" $-1 < g - i < 1$, using SDSS g and i filters . Fig. 3, shows the statistics for 18.5-19 mag bin, which illustrates the discrepancy : a nonzero SF seen even for the non-variable, blue stars. Quasars exhibit SF on the similar level to blue stars for short timescales, which means that their variability at those timescales is consistent with non-variable objects (stars). To illustrate the point, on Fig. 3 we plot a quantity $\chi = \Delta_m/\mathrm{error}$, shown on a grid of four different $\log \Delta_t$ versus four magnitude bins.

For each magnitude-$\log \Delta_t$ bin we calculate a set of properties: the median CRTS lightcurve error, the robust standard deviation of Δ_m, median CRTS lightcurve day-averaged magnitude, median CRTS lightcurve average error, and] the robust standard deviation of the $\chi = \Delta_m/\mathrm{error}$ distribution. For stars, which have no intrinsic variability, the width of χ distribution should be 1. Fig. 3 shows the χ distribution for a grid of magnitude-$\log \Delta_t$ bins - it deviates from 1 for all bins, and any deviation from that is due to unaccounted errors. For this reason we define the $f_{c,B}$ - the correction factor from blue stars, as the width of χ distribution for blue stars in the short timescale bin ($\log \Delta_t < 1.7$) (see Table 1 for values of χ for other bins). Applying thus derived correction factors to the overall sample shows that the structure function for the blue stars is brought to values allowed by the sky photometric variability (0.05 mag level, green dashed line on Fig. 4). Values of χ for various magnitude- bins is shown in Tab. 1 .

5 CONCLUSIONS

Implications: * information from correspondence with M. Graham * other recent findings that are based on short timescale CRTS variability - all would be called into question

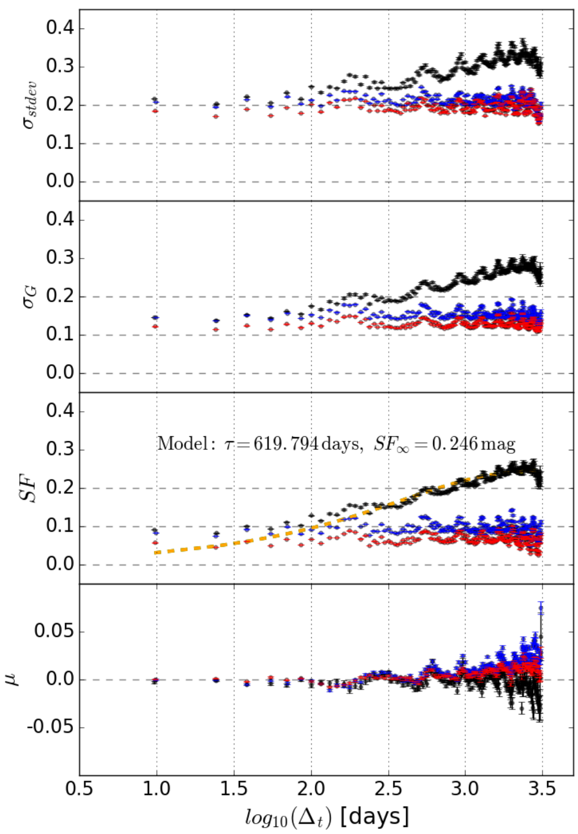

Figure 2. The four panels show statistics calculated for the subsample of 333 CRTS quasars (black points), 1400 "blue" stars (blue points), and 2087 "red" stars (red points), all chosen according to the SDSS r magnitude 18.5 < m < 19. For "red" stars we require that SDSS colors are $1 < g - i < 3$ and for "blue" stars $-1 < g - i < 1$. Lightcurve-derived pairwise brightness differences for all objects of a given type are binned according to linearly spaced 200 bins in Δ_t. The binning was not found to affect the main features of the plot. The sine-like modulation reflects differences in number of points in each bin (from tens to hundreds of thousands per bin). For each bin, we calculated for each type of object the standard deviation σ_{stdev}, the robust Gaussian standard deviation σ_G (from the interquartile range $0.7414(q_{75} - q_{25})$, SF and the mean(using eqs.5.67−5.68 in *Ivezic*+2004 (the AstroML book). Yellow dashed line on the SF panel traces the fiducial Damped Random Walk model. Note how both blue and red stars do not exhibit signs of variability as expected, whereas quasars (black) clearly show an intrinsic variability. At the low timescales $\log \tau < 1.7$ CRTS quasar SF departs from the fiducial model of Structure Function.

ACKNOWLEDGEMENTS

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Soci-

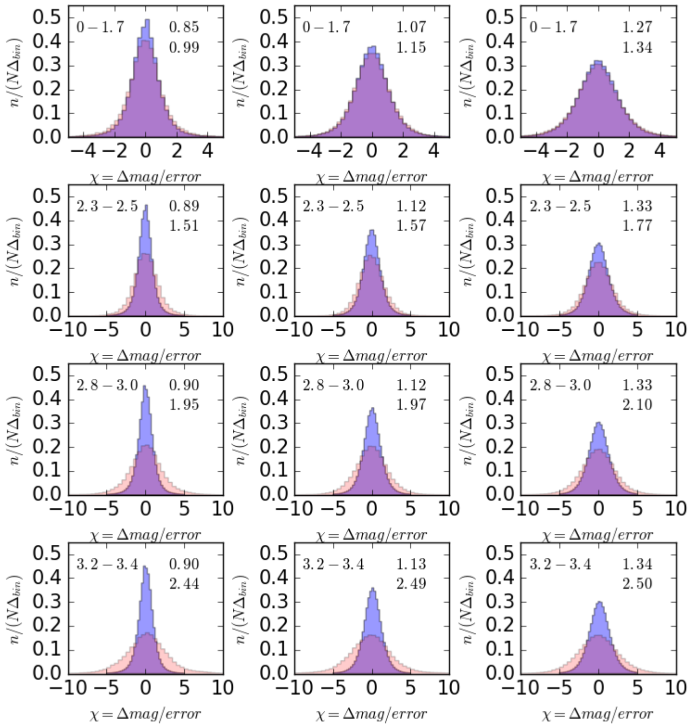

Figure 3. The top panels show how the histogram of $\chi = \Delta_m/\mathrm{error}$ for small timescales ($\log \Delta_t < 1.7$, $t < 50$ days) for quasars (red) overlaps almost perfectly with blue stars (blue). The implied small quasar variability is at the level measured by SDSS. On each plot numbers in the upper-left corner indicate the $\log \Delta_t$ range, and in the upper-right the robust width of stellar and quasar distributions of χ. From left to right, we iterate over the magnitude bins : 17 − 18, 18 − 18.5, and 18.5 − 19 mag. From top to bottom we change the $\log \Delta_t$ range. Note how the stars, being nonvariable, maintain the same spread of χ, due to their lack of intrinsic variability. The quasars spread more thanks to their intrinsic variability, but at small timescales their spread is the same as that of stars, consistent with their lack of short timescale variability.

Table 1. Values of χ for various magnitude- $\log \Delta_t$ bins for blue stars.

mag \ $\log \Delta_t$	0-1.7	2.3-2.5	2.8-3.0	3.2-3.4
17-18	0.864	0.907	0.916	0.915
18-18.5	1.091	1.143	1.150	1.160
18.5-19	1.298	1.359	1.360	1.368

ety, and the Higher Education Funding Council for England. The SDSS Web Site is http://www.sdss.org/.

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle As-

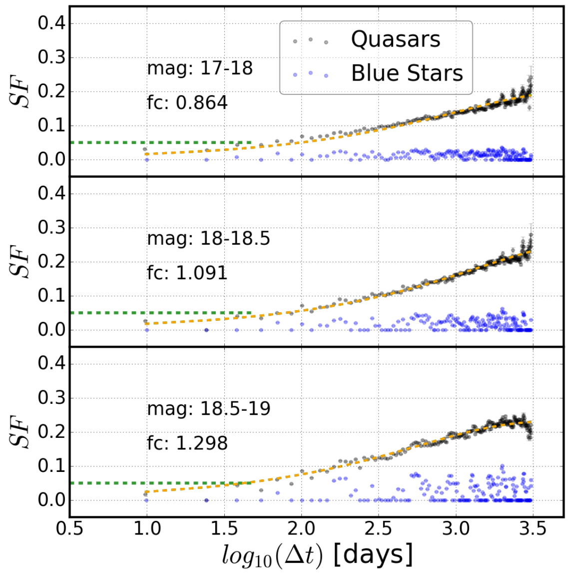

Figure 4. Three panels compare the Structure Function (SF) for quasars and blue stars in three magnitude bins based on their mean SDSS r magnitude. We correct the CRTS errors using listed fc factors. Stars have a flat SF consistent with their lack of variability, whereas quasars have a nonzero variability. At short time scales, the quasar variability seen in CRTS data is consistent with that from SDSS, shown by horizontal green lines.

trophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

This paper has been typeset from a TEX/LATEX file prepared by the author.