MNRAS 000, 1–4 (2015) Preprint 1 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.

Key words: keyword1 – keyword2 – keyword3

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 (Djorgovski+2012, Drake+2009) allowed an unprecedented study of well-calibrated light curves.

Reverberation Mapping studies (Peterson 2004, Fausnaugh+2016, ...) show that the Broad Line Region responds to stochastic variations in the continuum, meaning that the underlying variations originate from the accretion disk (Fausnaugh+2016, Munoz+2016).

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).

The AR(1) process (Damped Random Walk model) describes well the quasar variability at cadence longer than few days (Butler Bloom 2011, MacLeod+2010,2011,2012, Kelly+2009, Zu+2011,2013, Kozlowski+2010a, Ruan+2012), which corresponds to the observational capabilities of the CRTS survey.

[A deviation from a DRW at very short timescales, reported by Kasliwal+2015, is based on Kepler lightcurves at cadences shorter than a day, which is below the observa-

⋆ E-mail: suberlak@uw.edu

tional limit of CRTS. For this reason, in this paper DRW is assumed as a good description 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, which is well defined for the and specifically investigate the short-timescale regime of variability ($\tau < 54$ days) (MacLeod+2010, MacLeod+2012, Simonetti+1984, Vanden Berk+2004, Sumi+2005, Bauer+2009).

In this work we present the detailed error analysis of CRTS and SDSS data. In section 2 we describe the properties of our dataset, all selection criteria, and the structure function analysis. In section 3 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 4 and 5.

2 OBSERVATIONS

2.1 Data overview

The CRTS survey (Drake+2009) telescopes use 4kx4k CCDs, observations are made in unfiltered white light (Djorgovski+2011). As part of the CRTS pipeline this unfiltered photometry is calibrated to the V-band zero point. Our CRTS Quasar sample consists of 7754 Quasars, and after selecting only those that have more than 10 measurements, the sample is reduced to 7707 objects. Within that sample , 96% observations span the time of $7-9$ years. The light curve and sample-averaged mean and median error is 0.22. The mean and median brightness is 19.50 and 19.66 respectively. The mean number of individual observations per light curve is 209. 91.2% (7035 of 7707) quasars were observed between 1 to 4 times per night. (See Fig.2.1)

We complement the CRTS Quasars with information from position-matched SDSS Stripe 82 catalog[reference]. The catalog includes SDSS DR7 stripe 82 ($22h24m < R.A. < 04h08m$ and $|Dec| < 1.27deg$) spectroscopically confirmed 8,974 quasars cross-matched with 2MASS photometry.

As a control sample we use CRTS observations of standard stars in Stripe 82. Those are complemented with the SDSS photometry (Ivezic+2007) - we positionally match the CRTS observations to the SDSS calibration stars catalog (ver.2.6).

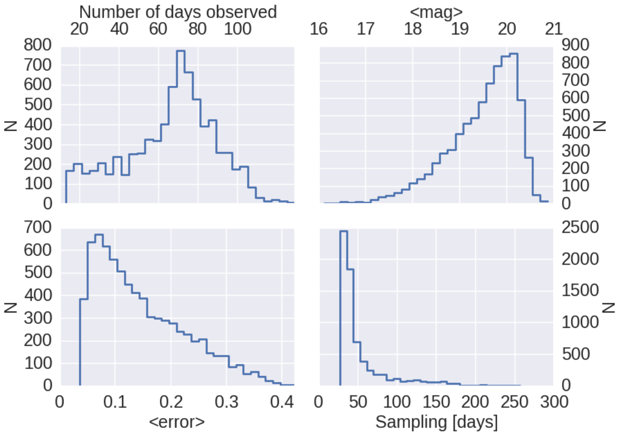

Figure 1. Statistical information about the CRTS Quasar sample including 7601 day-averaged light curves. The upper-left panel shows the number of days that a given quasar was observed, i.e. $max(MJD) - min(MJD)$ per light curve. The upper-right panel shows the average light curve magnitude. The bottom-left 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.

2.2 Day-averaging and selection

To improve signal-to-noise ratio CRTS light curves were day-averaged. Denoting individual CRTS reported measurement error as e_i, the combined error e_w consists of the sum of weighted errors where weight for each point is $w_i = 1/e_i^2$ (see eq.1).

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

If the resulting e_w was smaller than 0.02 *mag*, then 0.01 *mag* was added in quadrature. For further analysis we selected light curves that had more than 10 observation days, which reduced the sample to 7601 CRTS Quasars. A typical sampling interval of day-averaged CRTS Quasar lightcurves is ≈ 20 days.

2.3 Structure Function

The structure function 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 structure function 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: $e_{ij} = \sqrt{e_i^2 + e_j^2}$.

2.4 Sample Selection

We use CRTS standard stars as a reference for objects that exhibit no intrinsic variability (thus SF=0). To ensure a uniform sample we impose the following selection criteria, based on the SDSS r-band photometry: for CRTS quasars $17 < m_{SDSS,r} < 20$, and the CRTS lightcurve-averaged error to be $0.05 < \langle CRTS_{err} \rangle < 0.3$. Similarly, for the CRTS standard stars, $17 < m_{SDSS,r} < 20$, and the lightcurve-averaged error $0.05 < \langle CRTS_{err} \rangle < 0.3$. To assess how colors affect errors and variability, we divide the comparison stars into two color bins: "red" with $1 < g - i < 3$ and "blue" $-1 < g - i < 1$, using SDSS g and i filters .

To analyze the magnitude difference data we bin it in logarithmic time lag 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. Having performed tests with 50, 100, 200, 400 bins we find that 200 is the right choice between computational efficiency and preservation of scientific information. For each bin we calculate four statistics, shown on Fig. 2.4: Standard deviation (σ, Gaussian robust deviation from the mean $\sigma_G = 0.7414(q_{75} - q_{25})$, Structure Function ($SF$) and the mean ($\mu$). We calculate the SF using an exact prescription involving marginalizing the log-likelihood of the probability in $p(SF)$ and $p(\mu)$ space per bin, after [Ivezic+2013], chapter 5:

$$\left.\frac{dp(SF)}{dSF}\right|_{SF=mode} = 0 \qquad (2)$$

To show the departure of raw CRTS quasar points we fit to the SF a fiducial Damped Random Walk model:

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

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

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. If this was caused by an error underestimate, the variability on short timescales would be reduced by incorporating an error correction factor f_c. We find that simple increase of CRTS errors by 30% leads to a reduction of quasar SF to zero (see Fig. 2.4). This prompted a more detailed error analysis of points involved in short timescales on the plot, described in Sec.3.

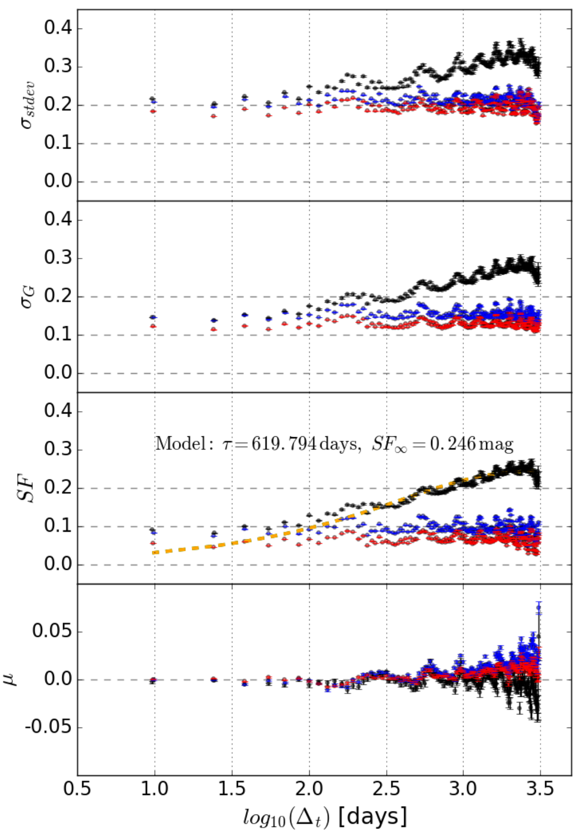

Figure 2. The four panels show statistics calculated for the subsample of 333 CRTS quasars (black points), 1400 "blue" stars (blue points), and 2087 "red" stars (red points), all chosen according to the SDSS r magnitude $18.5 < m < 19$. For "red" stars we require that SDSS colors are $1 < g - i < 3$ 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 $0.7414(q_{75} - q_{25})$, Structure Function 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.

3 RESULTS OF DETAILED SAMPLE ERROR ANALYSIS

To analyze the error properties of the short time lag points we select $\log \tau < 1.7$ points for CRTS quasars and stars, which includes 721283 time-lag points based on quasars, and 314248 time-lag stellar points. To disentangle the magnitude sensitivity, we further split the sample into three magnitude bins : $17 - 18$, $18 - 18.5$ and $18.5 - 19$ ^{mag}.

We separate contributions from the intrinsic variability and the error-related variance by comparing CRTS stars to quasars. Assuming that each δm data point originates from a homoscedastic Gaussian distribution, with a standard deviation σ_{com} (that includes all intrinsic variability σ and error-related $f_c * e_i$):

$$\sigma_{com} = \sqrt{\sigma^2 + (f_c \cdot e_i)^2} \qquad (4)$$

and the mean μ, then the distribution of all magnitude difference points in our sample is a sum of Gaussians:

$$\frac{1}{N} \cdot \sum_{i=0}^{N} \frac{1}{\sqrt{2\pi}\sigma_{com}} . e^{-(x_i - \mu)^2/(2\sigma_{com}^2)} \qquad (5)$$

Since standard stars do not vary on level above $0.02 mag$ [reference], it is reasonable to assume for stars $\sigma = 0$, and a distribution centered around $\mu = 0$. With these parameters we fit the ensemble of Gaussian distributions per magnitude bin to a histogram of "blue" stars to recover an error correc-

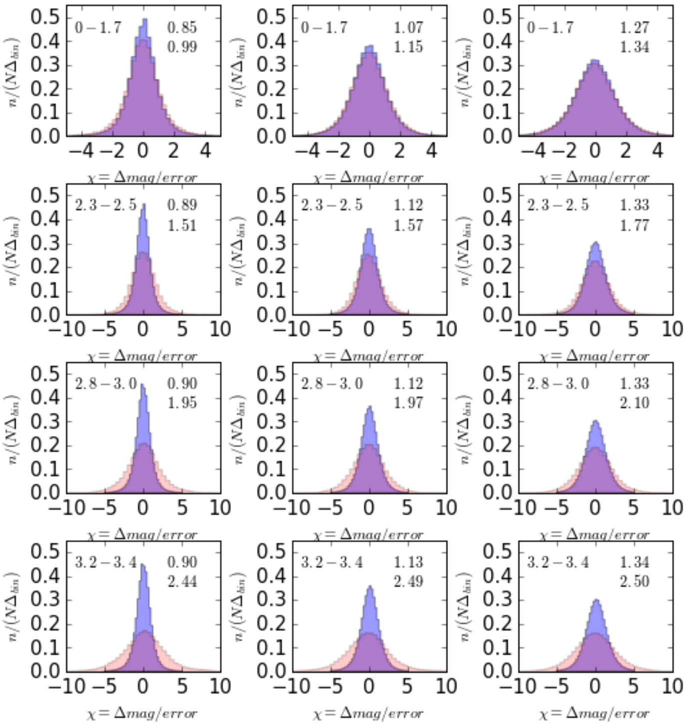

Figure 3. The top panels show how the histogram of $\chi = \Delta_m/\mathrm{error}$ for small timescales ($\log \Delta_t < 1.7$, $t < 50$ days) for quasars (red) overlaps almost perfectly with blue stars (blue). The implied small quasar variability is at the level measured by SDSS. From left to right, 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 same as that of stars, consistent with their lack of short timescale variability. 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.

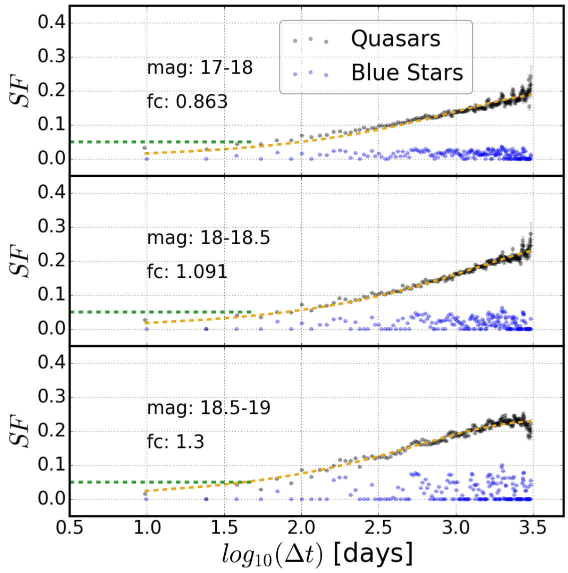

Figure 4. Three panels compare the Structure Function (SF) for quasars and blue stars in three bins based on their mean SDSS r magnitude. We correct CRTS errors using listed fc factors. Stars have a flat SF consistent with their lack of variability, whereas quasars have a nonzero variability. At short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

tion factor that would explain the variability ($\sigma_{com} = f_c \cdot e_i$. Those factors applied to quasar distribution explain their variability (see Figs. 3 and **??**).

4 DISCUSSION

5 CONCLUSIONS

CRTS errors were underestimated at the 20-30 % level.

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

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 Society, 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 Astrophysics 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.