MNRAS **000**, 1–1 (2015) Preprint 6 May 2016 Compiled using MNRAS LATEX style file v3.0

SDSS Stripe 82 : quasar variability from forced photometry

Krzysztof Suberlak,[1]⋆ Željko Ivezić, [1] Yusra AlSayyad,[1]
[1] *Department of Astronomy, University of Washington, Seattle, WA, United States*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

1 INTRODUCTION

Many objects in the universe, from stellar to extragalactic scales, vary on timescales less than a few hundred years. Lightcurves carry a wealth of information allowing one to infer various physical properties of a planet or a galaxy. If a lightcurve is poorly sampled, the inferred characteristics are less certain. Yet, since all astronomical observations suffer from a detection threshold, a very faint variable object in some epochs may be undetectable. Forced photometry rescues the information from very faint epochs by performing a measurement in all epochs in a location from the co-added images. An inherent challenge to such set of measurements is an intepretation of noise-dominated flux. To circumvent this problem many studies apply a magnitude cutoff few magnitudes above the detection limit, which reduces the amount of available data. Indeed, in order to fully utilize information present in time-domain surveys, such as Large Scale Synoptic Telescope, Palomar Transient Factory, or Sloan Digital Sky Survey, and properly characterize faint variable objects, we need to properly handle the faint flux measurements. A new methodology would allow an unbiased study of such faint variable objects, including quasars, RR Lyrae, Cepheids, and a wealth of other variable sources. With the advent of precision time-domain astronomy surveys it is crucial to apply the best possible faint forced photometry algorithms and thus make full use of the data.

2 VARIABILITY

Variations in object brightness have two main origins: an error-induced noise, and an intrinsic variability. A lightcurve consists of a set of N measurements of brightness x_i with errors e_i. In this analysis we assume that x_i are drawn from a Gaussian distribution $\mathcal{N}(\mu,\sigma)$, and that errors e_i are homoscedastic. We describe this distribution with two parameters : mean μ, and width σ. To increase efficiency, we employ a two-step approach after Ivezic+2014. First, we find approximate values of μ_0 and σ_0, and then we evaluate the full logarithm of the posterior pdf in the vicinity of the approximate solution. With a Bayesian approach, we find $\mu_{full}, \sigma_{full}$ by maximizing the posterior probability distribution function (pdf) of μ, σ given x_i and e_i : $p(\mu|x_i,\sigma_i)$ (see Fig. 2, and Appendix B for the detailed calculation).

For each lightcurve, we also calculate mean-based χ^2_{DOF} and median-based χ^2_R (the latter is more robust against any outliers in the distribution) :

$$\chi^2_{dof} = \frac{1}{N-1}\sum\left(\frac{x_i - < x_i >}{e_i}\right)^2 \quad (1)$$

and

$$\chi^2_R = 0.7414(Z_{75\%} - Z_{25\%}) \quad (2)$$

with $Z = (x_i - median(x_i))/e_i$.

Initially, we evaluate μ_{full}, σ_{full}, χ^2_{dof}, and χ^2_R for the entire lightcurve. Then, only if either $\sigma_{full} > 0$ or $\chi^2 > 1$, which hints some intrinsic variability, we also calculate μ_{full}, σ_{full}, and χ^2 for the seasonally-binned portions of the lightcurve.

On Fig. 2 we plot the stages of calculating μ and σ. Left and middle panels compare the two methods of calculating variability parameters. The initial approximation (dashed) is based on bootstrapped resampling of (x_i, e_i) points from the lightcurve. By randomly resampling the lightcurve M times, instead of a single sample with $N \approx 10-70$ points we have M samples. The histogram of $M = 1000$ values for μ, σ from resampling is plotted with dashed lines. We use the approximate values to provide bounds for the $200x70$ grid of μ, σ, used to evaluate the full posterior likelihood density function (right panel). This ensures that, despite using a coarse grid to improve computational speed, we still resolve the peak of the underlying distribution.

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

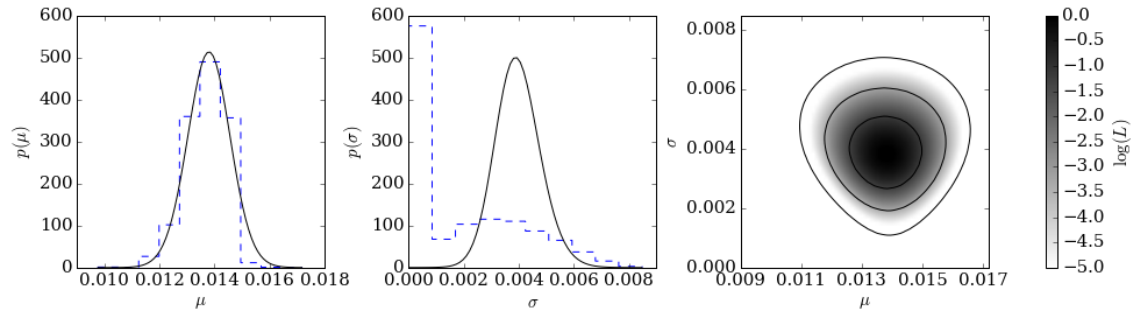

Figure 1. Two-step approach to finding μ and σ via μ_0 and σ_0 for an object 217720894888346446. In this calculation we use raw psf flux, before employing the faint source treatment outlined in Section **??**. On the left and middle panels, solid lines trace marginalized posterior pdfs for μ and σ , while dashed lines depict histogram distributions of 10,000 bootstrap resamples for μ_0 and σ_0. The right panel shows the logarithm of the posterior probability density function for μ and σ.