

# SDSS Stripe 82 : finding quasars in the forced photometry haystack

Krzysztof Suberlak,<sup>1\*</sup> Željko Ivezić,<sup>1</sup> Yusra AlSayyad<sup>1</sup>

<sup>1</sup>*Department of Astronomy, University of Washington, Seattle, WA, United States*

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

We provide variability and color-selected quasar sample up to  $g < 23.5$ . We reproduce many results from previous research, and find that using the LSST Stack reprocessed SDSS S82 data allows reaching fainter objects, and extending the reach of possible quasar candidate classification.

## 1 INTRODUCTION

This report aims to outline the process of analyzing the forced photometry reprocessed SDSS Stripe 82 data with the aim of improving the quasar selection using combined color and variability cuts.

Quasars are some of the brightest sources of radiation in the Universe. Most galaxies have gone through a phase of rapid accretion onto the central supermassive black hole, which resulted in emission of radiation from the hot accretion disk. As a result of physical processes in the disk, the radiation exhibits stochastic variability that can be mathematically described by a damped random walk, or a process with a certain covariance, and characteristic decay timescale. The amplitude of variability and characteristic timescale can be linked to physical properties of the disk, and are therefore of high interest. Beyond that, quasars are relevant as astrophysical probes, since the quasar luminosity function is related to the evolution of galactic initial mass function. (McGreer+2013)

Traditionally quasars can be found by color cuts because at least the nearby ones occupy a specific region in the color space (Sesar+2007, Ivezić+2003, Bovy+2011). However, it has long been recognized (Fan+1999) that as the redshift of the quasar increases, its color changes because we probe different regions of the intrinsic spectral energy distribution. Around redshift 2 quasars cross through the stellar locus in  $u-g$  vs  $g-r$  color space (Yang+2016, Richards+2015, Jiang+2014), so that without an additional selection criterion they are indistinguishable from stars (especially M dwarfs at redshifts 5-6, Yang+2016). At higher redshifts quasars again occupy a region that can be confused with RR Lyrae in the color space. Variability has been successfully employed for quasar selection for a number of years (Palanque-Delabrouille 2011, 2013, 2016, Schmidt+2010, VandenBerk+2004, MacLeod+2011, MacLeod+2013, Peters+2015). This is possible because stars in the main stellar locus are not variable, and RR Lyrae or Cepheids have a very specific variability pattern, very distinct from DRW. Eclipsing Binaries exhibit only very occasional deep dips in magnitude, which would be well distin-

guished from quasars also based on variability alone. Combining variability and color information can therefore provide a way to select quasars with minimally small contamination.

Stripe 82 is a very special SDSS field, observed multiple times, initially as part of the supernova survey. Each object in this equatorial stripe has between 60-180 epochs, and this allows study of variability. Coupled with supreme, well-calibrated SDSS photometry, S82 is a favorable testbed for selection and classification studies.

The S82 data was reprocessed in the Summer of 2013 as a result of the LSST Data Challenge, that was designed to test the capability of then in-development LSST Stack<sup>1</sup>. The S82 forced photometry dataset was also used as the testbed of database ingest into the Prototype Data Access Center<sup>2</sup>.

The reprocessing included preparing  $i$ -band coadds, source detection on the coadds, and forced photometry on these locations in all epochal  $u,g,r,i,z$  data. The reprocessing effort was shared between the NCSA and IN2P3 to test the portability of algorithms. A five degree overlap was processed in both Data Processing Centers (DPC).

## 2 S82 REPROCESSING

We use data from all SDSS runs up to and including run 7202 (Data Release 7), including all 6 SDSS camera columns. Stripe 82 survey covered an equatorial strip of the sky, defined by declination limits of  $\pm 1.27^\circ$ , extending from R.A.  $\approx 20^h(320^\circ)$  to R.A.  $\approx 4^h(55^\circ)$  (??). Observations conducted prior to September 2005 (part of SDSS I-II) had a more sparse sampling than SDSS-III, and the SDSS Supernova Survey, which ran between September 1st - November 30th each year between 2005-2007.

The SDSS Stripe 82 DR7 data was processed in two

<sup>1</sup> see <https://confluence.lsstcorp.org/display/DM/Properties+of+the+2013+SDSS+Stripe+82+reprocessing>

<sup>2</sup> DMTN-029 : Loading SDSS Stripe 82 data into PDAC <https://dmtn-029.lsst.io>

data centers : NCSA (National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, IL) and IN2P3 (Institut national de physique nucléaire et de physique des particules in Paris, France). NCSA processed data from  $-40\text{ deg} < RA < +10\text{ deg}$  and IN2P3 with  $+5\text{ deg} < RA < +55\text{ deg}$ . There is a  $5\text{ deg}$  overlap, used to confirm that the data processing pipeline in both data centers yields identical data products. The entire strip was split into smaller patches<sup>3</sup>

All epochs (individual images) were background-subtracted, and then scaled from the Digital Unit counts to fluxes by comparing standard objects against the (?) catalog (similar to Jiang+2014). We downloaded the data from the NCSA storage at <https://lsst-web.ncsa.illinois.edu/~yusra/S13Agg/rawDataFPSplit/>.

The available data includes the DeepSource tables, which constitute of i-band coadd detection information, and the DeepForcedPhot tables, which contain forced photometry seeded from i-band detections.

The DeepSource tables and DeepForcedPhot tables can be joined on deepSourceId, which is called objectId in DeepForcedPhot.

Yusra AlSayyad and Ian McGreer calculated summary aggregate metrics on the S82 S13 data, using the code available at <https://github.com/imcgreer/QLFz4>. This information was used in McGreer+2013, and AlSayyad 2016 PhD Thesis.

Source extraction often needs to address the problem of blended sources. In such case the process of deblending assigns a single ParentSourceId to a blended clump. Sources with parents brighter than 17 mag in the coadded i-band are considered unreliable, since they were not handled well by the deblender. ParentSourceId is null for objects that are their own parents (or in other words, are not blended).

Imposing a cut of 23.5 mag in coadded i-band, the main DeepSource NCSA -processed catalogs used contain 5474350 primary sources, and 1957486 non-primary sources (deblender parents and secondary detections). IN2P3-processed portion of the S82 includes 4998901 primary sources, and 1882303 non-primary sources.

Both NCSA and IN2P3-processed regions of S82 were divided into smaller sections called patches. There are 11 patches per each DPC :

The forced photometry tables, organized by filter-patch, contain the `id`, `objectId`, `exposure_id`, `mjd`, `psfFlux`, `psfFluxErr`, sorted by `objectId`, where Flux is measured in calibrated  $[ergs/cm^2/sec/Hz]$ .

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.

<sup>3</sup> patch boundaries can be found <https://lsst-web.ncsa.illinois.edu/lsstdata/dr-w2013/coaddBounds.txt>