

Time Domain Explorations With Digital Sky Surveys

Ashish A. Mahabal, S. G. Djorgovski, M. J. Graham, Priya Kollipara,
Benjamin Granett, Elisabeth Krause, Roy Williams, M. Bogosavljevic
California Institute of Technology, Pasadena, CA 91125

C. Baltay, D. Rabinowitz, A. Bauer, P. Andrews, N. Ellman, S. Duffau,
J. Jerke
Yale University, New Haven, CT 06520

A. Rengstorf, R. Brunner
NCSA/UIUC, Champaign, IL 61820

J. Musser, S. Mufson, M. Gebhard
Indiana University, Bloomington, IN 47405

Abstract. One of the new frontiers of astronomical research is the exploration of time variability on the sky at different wavelengths and flux levels. We have carried out a pilot project using DPOSS data to study strong variables and transients, and are now extending it to the new Palomar-QUEST synoptic sky survey. We report on our early findings and outline the methodology to be implemented in preparation for a real-time transient detection pipeline. In addition to large numbers of known types of highly variable sources (e.g., SNe, CVs, OVV QSOs, etc.), we expect to find numerous transients whose nature may be established by a rapid follow-up. Whereas we will make all detected variables publicly available through the web, we anticipate that email alerts would be issued in the real time for a subset of events deemed to be the most interesting. This real-time process entails many challenges, in an effort to maintain a high completeness while keeping the contamination low. We will utilize distributed Grid services developed by the GRIST project, and implement a variety of advanced statistical and machine learning techniques.

1. Introduction

Systematic exploration of previously poorly covered regions of the observable parameter space is a major source of discoveries in astronomy (Djorgovski et al. 2001a, 2001b). In particular, exploration of the time variability on the sky over a broad range of flux levels and wavelengths is rapidly becoming a new frontier of astronomical research (Paczynski 2000, Diercks 2001). All manner of variable stars, stellar explosions such as SNe and GRBs, variable AGN, pulsars, microlensing events, etc. are some of the examples of currently known exciting

time domain astrophysical phenomena. Many more are certain to be discovered over time as planned ambitious synoptic sky surveys (Tyson 2002, Kaiser 2002) join the existing ones involving the time domain (see, e.g., Paczynski 2001, <http://www.astro.princeton.edu/~bp/> for listings). The important factors in such programs are (1) the area covered, (2) the depth of coverage, (3) number of wavelengths used, and (4) the baseline(s) in time.

We briefly describe here a pilot project we carried out with the plate overlap regions of the Digital Palomar Observatory Sky Survey (hereafter DPOSS). We then outline ongoing programs with the Palomar-QUEST sky survey (hereafter PQ) and describe a realtime transient detection system and its challenges.

2. Variability with DPOSS overlap regions

DPOSS covers the entire Northern Hemisphere in three filters viz. J, F and N plates (calibrated to Gunn g , r , i). Each plate is 6.5 degrees wide with adjacent plates overlapping by strips that are 1.5 degrees wide. This results in at least 40% of the sky being imaged at least twice in each of the three filters. We conducted an exploratory search for highly variable objects (Granett et al. 2005) and optical transients (Mahabal et al. 2005) using $\sim 8000 \text{ deg}^2$ in the NGP and SGP areas of these plate overlap regions. The effective depth of these searches was $r_{max} \approx 19$ mag for the “high” states, with a plate limits $r_{max} \approx 21$ mag. Time baselines ranged from days to years, with ~ 2 yrs being typical. After eliminating various artifacts and contaminants, and applying well defined statistical criteria for selection, we identified a large number of highly variable objects, and followed up spectroscopically a subset of them at the Palomar 200-inch telescope. They turned out to be a heterogeneous collection of flaring M-dwarfs, OVV QSOs and BL Lacs, CVs (including a rare magnetic one), and some otherwise non-descript stars. Approximately a third to a half of all highly variable objects down to this magnitude level, at moderate and high Galactic latitudes appear to be associated with AGN.

We also found a number of optical transients (operationally defined as high-S/N, PSF-like objects, detected only once). We estimate that a single-epoch “snapshot” down to this flux level contains up to ~ 1000 transients/sky. Their nature remains unknown, but in at least 2 cases deep follow-up imaging revealed apparent faint host galaxies, which now await spectroscopy.

This pilot study gave us some hints as to what may be expected in a dedicated, wide-field, synoptic sky survey at comparable magnitudes. The faint variable sky has a very rich and diverse phenomenology.

3. The driftscans of Palomar-QUEST

The Palomar-Quest synoptic sky survey (Djorgovski et al. 2004, 2005; Baltay et al. 2005), a collaborative project between Yale, Caltech, and NCSA (some other groups are also involved in more specific collaborations) is a new major digital sky survey conducted at the Samuel Oschin 48-inch Schmidt telescope at Palomar. The survey uses a special 112-CCD, 162-Megapixel camera built especially for this purpose. Some of the salient features of the survey are: (1) Data taking in the Point-and Stare (PS) mode, covering $\sim 9.2 \text{ deg}^2$ per exposure,

or in a Drift Scan (DS) mode, in strips 4.6° wide, with a typical coverage of $\sim 500 \text{ deg}^2/\text{night}$, (2) Near simultaneous observations in one of two filter sets in the DS mode: Johnson-Cousins *UBRI* or SDSS *r'i'z'z'*, (3) In good conditions, typical limiting magnitudes for point sources: $R_{lim} \approx 22 \text{ mag}$, $I_{lim} \approx 21 \text{ mag}$, (4) In the DS mode, useful Declination range $-25^\circ < \delta < +30^\circ$, for a total anticipated survey area of $\sim 15,000 \text{ deg}^2$, (5) Multiple-pass coverage, with at least 4 passes per year at each covered location, (6) Time baselines for repeats ranging from days to months, anticipated to extend to multi-year time scales over the next 3 to 5 years or beyond, (7) NVO standards, protocols, and connections built in from the start.

The survey has started producing a steady stream of science-grade data, from summer of 2003. In the DS mode, it typically generates $\sim 1 \text{ TB}$ of raw image data per month (assuming ~ 14 clear nights). This unprecedented amount of data makes this the largest synoptic survey of its kind both in terms of area covered and depth. A broad range of science is envisioned for the survey, but exploration of the time domain will be one of the main focal areas.

PQ coverage as of Oct. 2004 is as follows: $\sim 11500 \text{ deg}^2$ have been covered in *UBRI* ($\sim 8900 \text{ deg}^2$ at least twice) and $\sim 12100 \text{ deg}^2$ in *r'i'z'z'* ($\sim 9300 \text{ deg}^2$ at least twice). We have been testing our transient detection techniques on areas that have been observed a large number of times. These techniques are being perfected in preparation of a realtime pipeline described in the next section. Besides the more exciting transient detections is also science which involves objects that simply vary in interesting ways. A large number of epochs with a range of baselines for several tens of thousands of quasars will provide good limits to differentiate between the models and further lead to better estimates of quasar lifetimes. This is possible by combining PQ data with earlier surveys like SDSS, DPOSS, DSS etc. Figure 1 shows B-band structure function for 500 quasars with $1 < z < 4$, a starting point for such a study.

4. Detecting Transients in Real Time

The existing *PQ* pipeline is geared to complete processing of a night's data by the next day. In a matter of hours catalogs become available in the four filters used and can be compared with other *PQ* epochs, or with other sky surveys and pointings for the area covered that night. This is sufficient for most *PQ* projects, including those involving variable objects, e.g. SNe. However, for the follow-up of rapidly fading sources and transients a faster turn-around time is needed. Thus, we have started work on a real-time processing pipeline which will enable detections of such sources within minutes or less. The following steps will be part of the process: (1) Compare nightly catalogs with older catalogs from *PQ* itself, (2) Compare the catalogs with other surveys and archives, using NVO infrastructure and methodology, (3) Compare positions of possible transients with those of known variable sources, known asteroids, etc. (see Figure 2 for an example), (4) A source classification engine will use the catalog data and the image data to categorize the possible transients to determine the likely types, (5) A decision engine will determine the significance of the object: should it be broadcast to the email subscribers list so that it can be followed up immediately, or should it go to the website to be announced as a lower priority but interesting

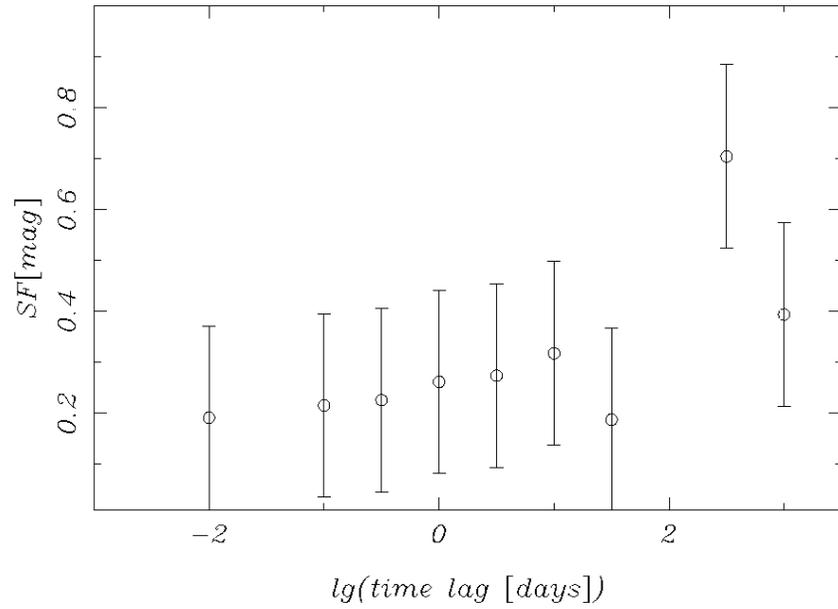


Figure 1. B-band structure function of ~ 500 quasars for $1 < z < 4$ using PQ and SDSS data. We are extending this to tens of thousands of quasars.

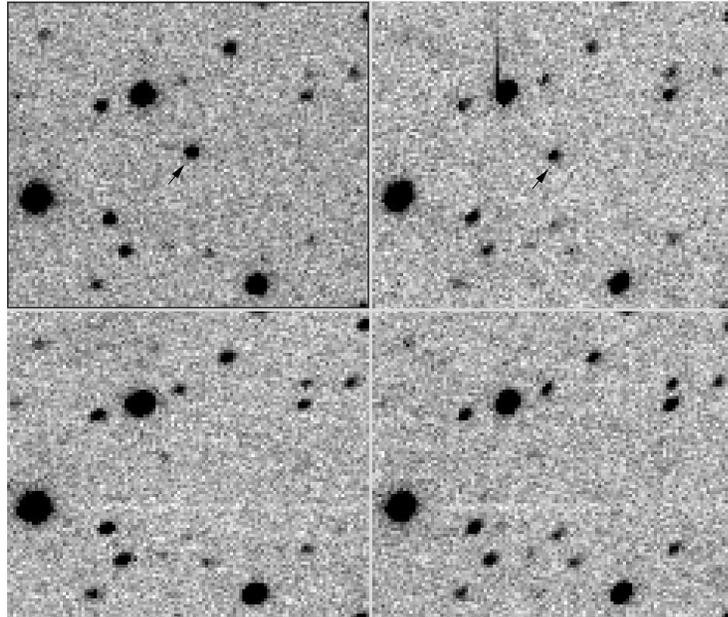


Figure 2. Example of an asteroid from R and I images from Aug and Jun 2004. The object easily passed as a star. Only after using the known asteroid database could it be identified with one. Faster asteroids are often easier to track as they form an elongated image.

transient, or just flagged as an interesting object to be looked at again in future epochs (in addition to other categories like possible variable, orphan GRB etc.)

Given the large area coverage of *PQ* and the results from our exploratory DPOSS project noted above, we estimate that we will be detecting up to several tens of highly variable or transient sources per night. Points (4) and (5) above summarize the key challenge of dealing with this abundance of data in an effective manner – maintaining a high completeness in terms of the interesting variable and transient sources discovered, while maintaining a low contamination rate by spurious or uninteresting sources – and doing it in real time. We will utilize distributed Grid based services developed by the GRIST project (Jacob et al. 2005) along with a variety of advanced statistical (Mahabal et al. 2004, Graham et al. 2005) and Machine Learning techniques towards this end.

Acknowledgments: This work was supported in part by the NSF grants AST-0326524, AST-0407448 and the NASA contract NAG5-9482. PK and EK were supported in part by SURF Fellowships at Caltech. SGD acknowledges a partial support from the Ajax Foundation.

References

- Baltay, C. et al. 2005, in preparation
- Diercks, A., 2001, in: Virtual Observatories of the Future, eds. R. Brunner, S.G. Djorgovski, & A. Szalay, ASPCS, 225, 52
- Djorgovski, S. G., et al. 2001a, in: Mining the Sky, eds. A. Banday et al., Berlin: Springer Verlag, p. 305
- Djorgovski, S. G. et al. 2001b, in: Astronomical Data Analysis, eds. J.-L. Starck & F. Murtagh, Proc. SPIE, 4477, 43
- Djorgovski, S. G. et al. (The PQ Team) 2004, BAAS, 36, 805
- Djorgovski, S. G. et al. 2005, in preparation
- Graham, M. J., et al. 2005, this volume, [P1.2.7]
- Granett, B., et al. 2005, in preparation
- Jacob, J., et al., 2005, this volume, [P1.3.8]
- Kaiser, N., et al. 2002, in: Survey and Other Telescope Technologies and Discoveries, eds. A. Tyson & S. Wolff, Proc. SPIE, 4836, 154
- Mahabal, A., et al. (The VOSTat Team) 2004, BAAS, 36, 806
- Mahabal, A., et al. 2005, in preparation
- Paczynski, B., 2000, PASP, 112, 1281
- Paczynski, B., 2001, in: Mining the Sky, eds. A. Banday et al., Berlin: Springer Verlag, p. 481
- Tyson, A., 2002, in: Survey and Other Telescope Technologies and Discoveries, eds. A. Tyson & S. Wolff, Proc. SPIE, 4836, 10