Time Domain Research in Astronomy

White Paper for the 2009-2010 Decadal Survey Committee

February 15, 2009

Author: Donald G. York, University of Chicago

Co-authors: Lifan Wang, Texas A&M University; Carl Pennypacker, Lawrence Berkeley Laboratories; Xiangqun Cui, Nanjing Institute of Astronomical Optics and Technology /NAOC; Don Lamb, University of Chicago; Alexei Khokhlov, University of Chicago; Michael Ashley, University of New South Wales; Stephan Basa, Laboratoire de Astrophysique de Marseille; Morley Blouke, Ball Aerospace and Technologies Corporation; Wei Cui, Purdue University; Geza Gyuk, Adler Planetarium and University of Chicago; Dan Holz (Los Alamos); Mark Hammergren (Adler Planetarium); Jon Lawrence, MacQuarie University/Anglo Australian Observatory; Roger Malina, Laboratoire de Astrophysique de Marseille; Anna Moore, Cal Tech; Reed Riddle, Thirty Meter Telescope Corporation; John Storey, University of New South Wales; Tony Travouillon, (Cal Tech); Brian Wilhite, Elmhurst College)

Abstract

We consider the need for a major new observatory to explore the time domain. Major areas of need include the determination of fundamental cosmological constants; the nature of gamma ray bursts; the physics and morphology of supernovae (SN Ia, SN II, impostor SN and intermediate luminosity transients); the frequency and nature of compact object binaries and mergers; the discovery of planets orbiting stars; and the probing of deep stellar interiors with asteroseismology. At the same time, a number of developments in instrumentation make possible a major push to finally cross into time domain astronomy in a statistically meaningful and multi waveband way. A major new facility capable of dedicated, large sky angle, continuous coverage on timescales of seconds to days, is needed.

I. Introduction

Just as sky limited capabilities have developed in all wavebands over the last few decades, so the capability to run time domain surveys, at various time scales, has increased. SWIFT and FERMI are leading the way in the high energy regime, SWIFT with an immediate X-ray and UV follow up that allow positions of gamma ray bursts to be determined. GALEX is running a survey mode operation in the UV for specific timescales. Surveys for SNe at a variety of redshifts (2-3 day time scales) have gone on for some years. The Allen telescope will eventually survey the entire radio sky frequently; PanSTARRS and LSST will provide large area sky coverage for objects with ten day timescales (it takes several observations to establish a transient). We discuss here the case of optical transient detection on short timescales with new facilities.

II. The technical problem.

Fast optical transient surveys have a long history, which we cannot review here. Basically, the results have been to find many false positives, and to create the knowledge base needed for more powerful instruments. The problem is three fold. First, very large sky coverage is needed, since the interesting objects are rare. Second, obtaining large sky coverage is very difficult because of all the natural interruptions: day/night, moon phase, weather, and equipment downtime. Finally, the critical follow up observations face many problems: for spectroscopy of gamma ray burst optical after glows, only 26% are followed up (Lamb 2007).

To quantify these concepts, note that the greatest continuous sky coverage now is with gamma ray observatories (SWIFT, FERMI, HETE and others). They achieve effective coverage of on order 10,000 square degrees of sky continuously, though not necessarily the same part of the sky all of the time. GRBs occur on order of once per day in the observable Universe. Therefore, assuming the bursts are one-time events, about 90 per year can be seen. The effective sky coverage is h, the product of the natural area of the telescope or array of telescopes that is continuously on the sky; the sensitivity of the equipment on the telescope; the efficiency of the software in sensing the target event types; and a set of terms that reduce the effective continuity of sky coverage (day/night cycles, moon phase, weather, seeing). The effective coverage must be evaluated for particular time scales of searched for transients, τ.

A given source type (cataclysmic variable, GRB optical afterglow, GRB orphan afterglow) can be estimated from models or from observing experience to occur with some frequency per square degree of sky per year. This rate is Φ. For GRBs, Φ = 0.01/sq.degree-yr.

The figure of merit for detecting particular transients on a given timescale is η=h x Φ, where η is the number of sources expected per year with equipment described by the parameter h, specified for given timescale, τ. LSST has h~2200 for searches for objects that rise and fall on timescales of τ~10^6 seconds (it takes at least three points to get information on a transient.) By comparison, ROTSE III (Rykoff et al. 2005), an array meant for follow up of GRBs, if used in a staring mode as a search instrument, has h=1.7 for τ=1800s to a limiting peak burst magnitude of 17th magnitude. (There is not space to review all of the transient detector experiments with repurposed, modern telescopes. An evaluation of h for different projects is given by York 2009 and York et al. 2006 (http://astro.uchicago.edu/~don/xian/). The only transient search projects with h > 2 are SDSS (York et al. 2000), used as a Type Ia supernova search (Frieman et al. 2008) on scales τ~10^6 sec to limiting magnitude ~22 (h=24), and arrays of small telescopes with limiting V magnitudes < 14th (h < 1000) (see, e. g., Burd et al. 2005, NewA, 10, 409) which can detect only a few extragalactic optical transients). Note that SDSS, used to search for variable objects on timescales of 60 seconds by comparing different filters, has h = 0.01. LSST, repurposed to search for transients with τ=30 sec, has h = 0.6.

Some well known, short, extragalactic transients with timescales < 10^6 sec (12 days) have values of Φ (to limiting magnitudes given in parentheses) as follows: GRB afterglows (14-23), Φ = 0.01 – 0.02; orphan afterglows (23), Φ > 2; the blue flash of Type II SNe (21.5), Φ = 0.04; brightening from lensed QSOs (22), Φ = 0.01; the aggregate of SNe (23), Φ = 4.4; AGN (21), Φ= 100 (they vary on all time scales, all the time). For some well-known Galactic objects, AM CVn stars (19), Φ ~ 0.001; eclipsing planets and lensing planets (22), Φ ~0.01; cataclysmic variables (21) and red dwarf flare stars (23) , Φ = 1.

It is thus clear that the listed transients, except for AGNs, require observing systems with h >> 100 to make numbers of detections approaching a statistical sample (>>1). The equipment applied so far to faint optical transit searches have h ~ 1. It is obvious why existing systems have not found large numbers of extragalactic transients.

For the purposes of this paper, we consider a survey for short transients with h ~2000 on all timescales τ < 10^6 seconds. To carry out the scientific surveys required by the science presented below, the requirements for such a project would be to 1) achieve point source detections in 10 seconds to an effective limit of 20th magnitude in V; 2) maximize h by choice of site and continuous, large sky angle coverage; 3) build software to automatically detect transients on all time scales and to reject false positives automatically and instantly (see the appendix); 4) provide equipment on-site to follow up those transients with timescales < 1 day; 5) provide an alert system to trigger on-site follow-up equipment and to notify observatories world wide of all transient detections, on all timescales; 6) design all equipment to be robust, to minimize operations costs; 7) provide an ability to do surveys in different bands or in polarization; 8) built a human network such as Galaxy Zoo to enable very large numbers of objects to receive follow-up.

III. Science cases

A. Supernovae

The specifications for an h=2000 survey follow from the need to understand supernovae. The details of supernovae explosions are still a mystery because of the variety of mechanisms that may be involved, particularly in hypernovae and Type Ia SNe, and environmental issues. There is the possibility, on the one hand, that dust, formed in the rapid expansion, may affect the observations (Wang 2005). On the other hand, there may be additional bodies in the proximity of the SNe that may be affected by the blast wave and cause extensions of the light curve. High frequency variability observations may shed light on the detailed physics of the explosions.

For Type Ic supernovae, there is a tie-in with gamma ray bursts (GRBs) (Stanek et al. 2003), which appear a few days after some GRBs (further discussed below).

For Type II SNe, the crossing time is tens of minutes and the blast wave may be observed at the surface, creating a precursor event and allowing the light curve to be followed over the earliest hours of the SNe. For Type I SNe, we really have little idea of what the first few minutes might look like, but early activity may be observable in the brightest objects. The apparent luminosity of GRBs is so high that they must be beamed, but there are limits consistent with relativistic plasma being unassociated with most SNe (Gal-Yam et al. 2006). On the other hand, the jet picture for the rare Type Ic/GRBs (Lamb et al. 2005) must be only a first approximation, because the optical and X ray afterglows do not behave in a purely geometrical way (Panaitescu et al. 2006). It is not clear why the afterglows are so extended (up to 30 hours, Gomboc et al. 2005) and have structure. Some are not visible in the optical, perhaps because of dust in or near the object.

A persistent theme of the GRB data for long GRBs is that either there is material exterior to the original source, interacting with the jet, or there is something that keeps the central engine going for an extended period, such as infall after the burst. The short, hard bursts are defined as being < 2 sec in duration, but Norris and Bonnell 2006 found a one minute tail when many BATSE short bursts were summed, so even the classifcation of the GRBs is still uncertain. A strong test of the jet paradigm can be made if off-axis afterglows (orphans) can be seen. For current models, the predicted numbers are comfortably high for confirmation (Zou et al. 2007, Rossi et al. 2009) with an h=2000 transient survey, but sampling on time scale of hours is necessary to safely define the light curves. There are numerous theories that have been invoked to explain the details of these objects, but the observational situation needs to be better defined. Ground based detection of the GRB afterglows, without the GRB trigger, will allow precursors to be identified (Cenko et al. 2006) and the obtaining of extensive follow-up observations, to explore the presence of dust in the earliest phases, to measure polarization while the source is very bright and to obtain spectra throughout the event, is crucial: GRB research is in the early stages. Simultaneous optical, IR, gamma ray and X ray observations will be available while SWIFT is alive and, in gamma rays for sometime to come, with FERMI.

A near-coincidence with neutrino bursts (Razzague et al. 2004) or gravity wave signals (Abbott et al. 2006) is expected. Initially, the association may need to be established by co-adding signals in neutrino and GW data streams at given the optical and gamma ray bursts. Eventually, neutrino detectors (ANTARES, ICE CUBE) are expected to detect bursts of neutrinos from the GRBs. As noted later for (larger) stellar mass black holes, optical flashes in precursor or afterglow mode, with GRB and neutrino signals (which may not be time coincident) will form a critical empirical base for understanding GRBs (Basa et al. 2008).

B. Cosmology and Intergalactic Medium

Distance Scale. A transient detector such as Xian will make major contributions to the determination of cosmological constants and the probing of the high z Universe. The mean redshift of the GRBs is 2.8 (Jakobsson et al. 2006) and they may be observable at z>6. Lamb (2007) as reviewed the prospects of probing the Dark Ages using GRBs. The redshifts can be obtained from absorption lines local to the GRB (Prochaska et al. 2006). The same data can be used to examine gas abundances in the host galaxies and in intervening galaxies. At high z, these observations complement those in the very rare high z QSOs. The follow up must be done on-site (see the previous note about the incompleteness of current followup of this type.

Hubble constant. The use of variable multiple objects of lensed QSOs by foreground galaxies (Kundic et al. 1997), to determine H_o, will be possible. Low amplification lenses are more frequent than the better known cases with amplifications of 50 (see Inada et al. 2009) and for these, the timescales for variations between different lensed images is days to months. These can be found in an h = 2000 survey. Supernovae will be multiply lensed by cluster potential wells, complementary to determination of H_o from QSOs (Bolton and Burles 2003). The timescales are months to years but continuous tracking is crucial to see the flux changes in one image of the lens mirrored in the other. Once the variations are found, the objects can be monitored from other observatories.

Dark Energy. The number of lenses over the sky is a measure of the amount of dark energy in the Universe (Oguri et al. 2008). These can be found by variability searches. The use of standard Type Ia supernovae to determine the dark energy content is well known (Clocchiatti et al. 2006). Type Ia supernovae will be seen within one day of the explosions for thousands of objects (from above, $\eta = h \times \Phi = 2000 \times 4.4$) each year and light curves of one hour resolution will be available. Even if the h = 2000 survey instrument is in a white light mode, the data following detection can be used to bridge gaps in the follow-up observations that should generally occur off site.

Black hole formation and AGN outflows. Stellar mass black hole formation may be indicated to be occurring in young, star forming clusters (Heike et al. 2003). These will trigger X-Ray transients, but large scale X ray transient surveys are not possible in the near future. SWIFT XR and UVOT refinements of gamma ray positions allowed follow up. So, optical transient detection from forming black holes could trigger conclusive monitoring programs of particular sources in high energy bands (HESS, VERITAS, AUGER, FERMI), or lead to programs pointing to ways to better utilize the archives of these instruments for prior activity. Similarly, time dependent optical signals from AGNs may provide optical precursors whose detection could trigger timely use of the high energy instruments, or their archives (Poggiani, 2006; Chatterjee et al. 2008).

Interactions of various objects with BHs will be picked up with the advent of gravitational wave (GW) astronomy, first with Advanced LIGO, then LISA. These may be implicated in some of the Chandra detections noted above. An aggressive program of optical detections to match with GW events in position and or time coincidence for the always-on, all-sky GW detectors, will be essential to understanding the origin of the GWs. In particular, optical bursts could provide the critical coordinates that lead to redshift determinations.

C. Variable Stars

Cataclysmic variables (CVs: Rau et al. 2007) and dwarf flare stars (Kulkarni and Rau, 2006) provide the largest contamination sample in a survey of extragalactic transients: their systematic identification will make the extragalactic transient survey more effective and will qualitatively change our understanding of the two stellar types.

Cataclysmic variables consist of compact objects accreting mass from lower mass companions. The periods range from 10 minutes to hundreds of days. In a transient survey to >23rd magnitude over a day, dwarf novae will be the most frequent false positives (duration 3-5 days, periods of 40 days, burst magnitude ~5 magnitudes). The proposed survey should find virtually all of these in the Galaxy, out of the disk, and they can be seen in their quiescent state in a $V_{\rm lim} = 25$ all-sky template made with the h = 2000 survey itself (in one day). (See appendix.)

There are numerous manifestations of CVs, scientifically interesting because of the various mechanisms at work in their formation and evolution: mass transfer of deep layers of the secondary, magnetic braking, and gravitational radiation, the latter most manifest in the rare AM CVn stars (Anderson et al. 2005), likely to be seen by LIGO and candidates to become Type Ia SNe. A transient survey that is complete would allow a tie in between a well-defined CV sample and surveys with strong selection effects but of a form that reveals the physics of selected objects in great detail (Szkody et al. 2006). X-novae are similar in some ways to CVs and they flare on the same wide range of time scales, and are seen in all many energy bands, so the optical detections will serve as triggers for high energy observations in that case as well.

Dwarf flare stars (M stars that emit bursts of UV radiation) will similarly be eliminated by false positives. A large catalogue of such stars will, however, result. They can be used to understand the magnetic fields in late M stars and to understand the habitable zones for planets around such stars. Planets may be easier to find than in hotter main sequence stars, but the lower quiescent optical/IR energy and the flare impact on proximate planets may severely limit the life forms that can exist in such systems.

Asterioseismology to probe the stellar interiors of stars across the HR diagram will be possible for objects with periods of minutes to hours to days that cannot be done well where the day/night cycle or clouds interrupt observations (Winget 1998, Saio et al. 2006). Irregular variables from pre-supernova stars to forming stellar accretion disks will also obtain unprecedented coverage because of their erratic nature and the continuous nature of the survey data streams.

Finally, low mass planets can be observed in lensing situations (Beaulieu et al. 2006) and in eclipsing situations. The former case involves using machos for which the foreground stars have planets. Most studies are concentrating on the bulge of the Galaxy since the number of Galactic objects outside the bulge is small (Han 2008, see Table 1); and all sky survey would be free of selection effects. Eclipsing planets can be studied by the attenuation of the stellar light by a transiting planet on time scales of hours, for the brightest objects that can be observed without saturating the CCDs. The eclipsing systems are critical for the study of planetary atmospheres by high resolution spectroscopy, and help obtain estimates of the planet age, composition and surface temperature. To some limiting magnitude, depending details of the performance of the proposed survey equipment at high counts, most of these in the sky survey area will be found, since they are periodic and many periods will be observed.

Conclusion

A number of cases are presented for which the lack of time dependent information on short scales deprives astronomers of descriptive information needed for classification of objects at the most fundamental level. For lack of such information, a number of theoretical descriptions of some of the most interesting objects in astronomy are ambiguous or incomplete. Technology now makes it possible to build a very large optical survey instrument for short time scale monitoring of transients ($<10^6$ seconds). Examples of extragalactic and galactic problems are presented to show the need and the impact on astronomy that such an instrument will have.

Appendix

Advances in CCDs with multiple readout preamplifiers, in design of telescope optics and in computer real-time and data distribution techniques make an h~2000 optical survey for short transients possible at this time. A reasonable compromise to reach 20th magnitude with an array of small telescopes (to cover large sky angles) is to accept a 10 sec limitation on τ (rough values, of course, depending on details of implementation). That limit and seeing of one arcsec is acceptable for transient detection for the science programs mentioned.

Such a survey cannot be done, however, unless all the false positives are accounted for. For a survey for extragalactic variables, the main background of false positives (Φ in parentheses) is due to cataclysmic variables (Φ =1 to 19th magnitude), red dwarf flares (1 to 23rd magnitude), asteroids (1000 to 23rd for ROTSE, 1,000,000 for SDSS), cosmics (10 million). The asteroid estimate is for those that are not fast movers.

The detection software must take account of lessons learned from past surveys, which implies having a continuous template of the sky (Djorgovsky et al. 2008; York et al. 2009), remade at least every three months or so, on the basis of the data from the survey itself. If it needs to be made more frequently, the data to do that requires 24 hours of data of the project itself, so the only issue is computer processing power. The template must be 5 magnitudes deeper than the transient survey limit.

The detector system can then rely on drift scan mode, and on comparison, in the first instance, of each line read out with the previous line at that pre-amplifier (for rejection of moving objects and many cosmics) and on comparison of each read line with a constantly evolving local template based on the deep template noted above. The template must be modified dynamically to account for seeing local to a direction and observation, to account for predicted positions of slow moving asteroids and to account for attenuation due to clouds. The computer hardware to do this requires a high-end desktop computer for every 2×10^8 sky pixels, one arcsecond per pixel.

References:

Abbott, B., Abbott, R., Adhikari, R. 2006, PhRvD, 73, 2001.
Anderson, S. F., Haggard, D., Homer, L. et al. 2005, AJ, 130, 2230.
Basa, S. et al. 2009, Proceedings of the International Worksho on a Very Large Volume Neutrino Telescope for the Mediteranean Sea (2008)
Beaulieu, J.-P., Bennett, D. P. et al. 2006, Nature, 439, 437.
Berger, et al. 2009, ApJ, in press. (arXiv:0901.0710).
Bolton, A. S. and Burles, S. 2003, ApJ, 592, 17.
Burd, A., Cwiok, M., Czyrkowski, H. et al. 2005, NewA, 10, 409.
Cenko, S. B., Kasliwal, M., Harrison, F. A. 2006, ApJ, 652, 490.
Chatterjee, R. et al., 2008, ApJ 689:79-94.
Clocchiatti, A., Schmidt, B. P. et al. 2006, ApJ, 642, 1.
Djorgovski, S. G., Baltay, C. et al. 2008, AN, 329, 263.
Frieman, J. A., Bassett, B., Becker, A. et al. 2008, AJ, 135, 338.
Gal-Yam, A. et al., 2006, ApJ,639, 331.
Gomboc, A., Guidorzi, C., Mundell, C. G. et al. 2005, arXiv/0612762.
Han, Cheongho 2008, ApJ, in press.
Inada, N., Oguri, M., et al. 2009, ApJ, in press. (arXiv0809.0912).
Jakobsson, P., Levan, A. et al. 2006, A&A, 447, 897.
Keeton, C. R., Burles, S., Schechter, P. L. et al. 2006, ApJ, 639, 1.
Kulkarni, S. R. and Rau, A. 2006, ApJL, 644, 63.
Kundic, T., Turner, E. L., Colley, W. N., et al. 1997, ApJ, 482, 75.
Lamb, D. Q. et al. 2005, ApJ, 620, 355.
Lamb, D. Q. 2007, Phil. Trans. R. Soc. A, 365, 1363.
Norris, J. P. and Bonnell, J. T. 2006, ApJ, 643, 266.
Oguri, M., Inada, N., Strauss, M. A. et al. 2008, AJ, 135, 512.
Panaitescu, A., Meszaros, P. et al. 2006, MNRAS, 369, 2059.
Poggiani, R., 2006, ASpSci, 306, 17.
Prochaska, J. X., Chen, H.-W., and Bloom, J. S. 2006, ApJ, 648, 95.
Rau, A., Schwarz, R., Kulkarni, S. R. et al. 2007, ApJ, 664, 474.
Razzaque, S., Meszaros, P. and Waxman, E. 2004 PhRvL, 93, 181101.
Rossi, E. M., Perna, R., Daigne, F. 2008, MNRAS, 390, 675.
Rykoff, E. S., Aharonian, F., et al. 2005, ApJ, 631, 1032.
Saio, H., Kuschnig, R., et al. 2006, ApJ, 650, 1111.
Stanek, K. Z., Matheson, T. et al. 2003, ApJ, 591, L17.
Szkody, P., Henden, A., Agueros, M. et al. 2006, AJ, 131, 973.
Wang, L. 2005, ApJ, 635, L33.
Winget, D. E. 1998, JPCM, 10, 11247.
York, D. G., Wang, L., Cui, X. et al. 2006, SPIE, 6267, 46.
York, D. G., Adelman, J. et al. 2000, AJ, 120, 1579.
Zou, Y. C., Wu, X. F. and Dai, Z. G. 2007. A&A, 461, 115.