Large Synoptic Surveys

Lynne Jones
University of Washington

What is a "large synoptic survey"? While exact definitions of what makes a survey "synoptic" differ, the general characteristics of "large synoptic surveys" include: a sky sampling rate that is sensitive to astronomical phenomena that change over time, a large amount of sky coverage to provide a statistically significant sample of those phenomena, and data processing that can expose the transient or variable nature of the astronomical objects within the survey. Generally, these surveys are also intended to provide data on a wide variety of objects, leveraging the wide sky coverage and time sampling of the survey to serve many science goals.

The precise science goals of a particular survey depend on the scientific questions the survey planners wish to address. What is required are observations sensitive to time-varying phenomena - typically the variability is in the brightness of the object being studied but the position or color of the object could vary instead or as well. In theory, a record of the entire sky with very fine time resolution, high sensitivity (faint limiting magnitude), excellent astrometric and photometric accuracy, and multi-wavelength observations would yield enough data to characterize and study in-depth every kind of transient and variable object, requiring no or very few limits on these scientific goals. In practice, the goals of the survey have to be balanced with the resources available, with limitations on sky coverage, observational cadence (time resolution), sensitivity, wavelength coverage, and (to a lesser extent) accuracy and data processing capability.

How to achieve this balance and connect goals with observing strategy? This, of course, is a tough question with more than one potential answer. The final goal of the survey should be to provide a statistically significant sample (or samples) of the kind of object (or objects) of interest to the survey planners, with well-understood selection biases and completeness functions, and with as much data as necessary or possible for characterizing each individual member of each population, so that the underlying science question (or questions!) can be answered (e.g. "What does the orbital distribution of Trans-Neptunian objects (TNOs) imply about the evolution of the Solar System" "How do the distance and redshift distributions of Type Ia SN constrain models of the expansion of the universe?").

Achieving a statistically significant sample to address a particular problem implies that there is a number of target objects of a particular nature (whether this is at a particular redshift or spatial position or metallicity or just of a general classification) that must be discovered. As a first step, the luminosity function or number density function combined with an expected magnitude and distance distribution gives an estimate of the number of target objects which can be expected in a given area of sky coverage to a particular sensitivity limit (limiting magnitude).

Generally speaking, for a given telescope system there are tradeoffs between sky coverage and limiting magnitude: in a given total amount of time, either fewer fields can be observed with longer exposure times (and thus fainter limiting magnitudes), or more fields can be observed with shorter exposure times (and thus brighter limiting magnitudes). There are, however, some limitations: exposure times should not get much shorter than the readout plus slew/settle time of the telescope to avoid inefficient observing and should not be longer than the timescale of variability of the target object to maximize sensitivity to that variability. Within this envelope, there will usually be a range of acceptable exposure times, where the final trade between sky coverage and limiting magnitude will be set by the requirements of the science case. For example, if target sample size is most important and the luminosity function increases quickly with limiting magnitude, longer exposure times will be desirable (this is illustrated in Figure 1). If on the other hand, sensitivity to changes in the target distribution with location are important, greater sky coverage will be most desirable. It is possible to improve the final sample size result by going to a telescope system with a larger field of view or higher sensitivity due to a larger diameter primary mirror, a higher system throughput (due to better optical components, a more sensitive detector, or smaller atmospheric absorption) or a better point-spread function (better seeing), but the basic tradeoff is still present.

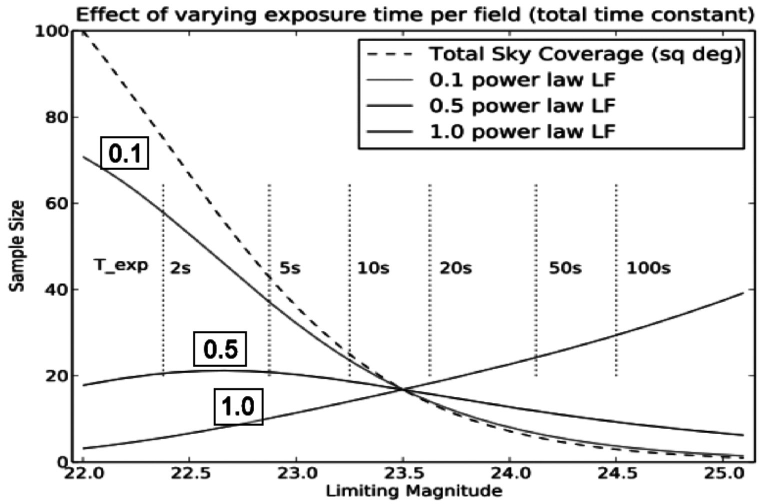

Figure 1: For a survey with a given total time span and throughput, there are tradeoffs between limiting magnitude in a given exposure and total survey area (see the black dashed line), that are parameterized by the exposure time. For a particular limiting magnitude and sky coverage (at a particular exposure time), the resulting sample size can be calculated. The sample size as a function of exposure time (indicated by the vertical dotted lines) for populations with power law luminosity functions of various alpha indexes (LF proportional to 10^alpha) are illustrated. Steep luminosity functions (1.0 line) produce larger sample sizes when less sky is covered to a fainter limiting magnitude; shallow luminosity functions (0.1 line) do better with a shorter exposure time and more sky coverage.

For transient and variable objects there is an additional complication to estimating the resulting sample size, which is the frequency, timescale and amplitude of the target object's variability. This variability must be detected to identify an object as a member of the target sample. The frequency of variability can be included into the estimated sample size at least roughly by just multiplying by the expected rate or duty cycle times the total survey lifetime. The effects of the timescale of variability are trickier, as there are vastly different time periods available for detecting different kinds of variable or transient objects. The optical counterparts of gamma ray bursts brighten quickly within minutes to hours then fade over hours to days, while supernovae brighten over a period of hours to days and then fade over a period of tens to hundreds of days; neither of these events repeat. M-dwarf flares occur and disappear within minutes but recur often although unpredictably, while periodic variables may vary on a timescale between hours to days but repeat on a definite schedule, making it possible to phase observations from different periods back together. Other objects display primarily astrometric variability rather than photometric variability - objects in our solar system display motions visible on timescales between seconds to hours (which may also carry them out of the survey area if the sky coverage is not large enough) while stars display parallax and proper motion changes which may only be visible on timescales of years. To detect any of these kinds of variability, multiple observations of the same object have to take place within the timescale of variability and often with a particular cadence within that timescale.

The photometric and astrometric accuracy of the survey also influence how often observations must take place in order to detect variability, as does the amplitude of the variability itself. Higher signal to noise measurements of the variability (due either to greater accuracy in the measurement or larger variability) mean that fewer observations are necessary before the variability can be detected and identified. In practice, accuracy is typically dominated by the effects of the limiting magnitude of the system itself and the amplitude and timescale of variability are determined by the target objects themselves. The major survey planning decision, especially for large synoptic surveys with multiple science goals, is how to reconcile the observing cadence required by each science goal with maintaining the required sample sizes, when sky coverage or exposure time (and thus limiting magnitude) must be reduced to allow time for multiple observations of the same targets.

However, there is yet another twist - how to identify the target objects of interest among the background of millions of other potential variable or transient objects? The observations described above are the minimum requirement to detect objects which are consistent with being members of the target population, but this results in a pool of candidates contaminated with other variables or transients. Reducing that contamination requires additional observational data either to determine the full shape of the transient or variable lightcurve, or to determine the velocity and acceleration vectors of a solar system object, or to determine the spectral energy distribution of the object through observations at multiple wavelengths or in multiple filters. For example, accurate identification of Type Ia supernova from their light curves requires multi-color observations of each potential SN separated by a few days, while identification of a fast-moving Near Earth Object (NEO) requires at least two observations, preferably in the same filter, with a separation of 30 minutes to an hour and repeated over several days within a period of 15 to 30 days, during which time the NEO has crossed a large portion of sky - requiring large sky coverage as well as frequent observations. The process of winnowing out the contaminants and identifying the true members of the target sample is the heart of "classification", which is described in more detail in section 1.3, but a primary requirement will always be "more data".

Luckily the required additional observations can come from many places. Often, and particularly for large synoptic surveys, it is most practical to generate at least some of these followup observations within the survey itself, translating to further balancing of survey time between sky coverage, limiting magnitude, and repeat observations of the same fields. For a target sample which is very numerous, with a high contamination rate but some potentially distinguishing characteristic (such as color or a lightcurve which is distinctive with additional data points), obtaining additional observations within the survey itself will likely be the most efficient method of reducing contamination and obtaining a better sample. Still, this is not always practical or even possible. Some kinds of objects are only identifiable with observations at wavelengths other than those achievable at the survey telescope, while others may need so many observations or different kinds of observations (such as spectra) that the survey telescope would be too far diverted from its primary purpose. In these cases, it is better to generate the followup data at other telescopes. Much of the rest of this book is devoted to the nuts and bolts of how to request this followup data through VOEvents and the process of communicating this information through the network of telescopes and observers available to obtain these additional observations.

A final consideration when planning a synoptic survey is how to process the data to detect the target objects in the survey. Briefly summarized, methods for detection of variables or transients boil down to generating a catalog of objects which are 'different' (in brightness or position) in an image acquired from the survey telescope from what is expected for that region of sky. One common method is to create a 'template image' from a series of images previously obtained of the same field, match this template to the image from the telescope being searched for variables or transients, and then to subtract the two to create a difference image. The matching process must include matching the point-spread function of the template and science image, as well as aligning their coordinate systems. Source detection algorithms are then run on the difference image to look for objects which varied in brightness or position between the template and science image. Another common method is to create calibrated catalogs from each image directly, match objects from the catalogs created from each image with the same objects from the other catalogs using spatial correlations, and again look for things which vary in brightness or position. Each method has its advantages and disadvantages, among which are: difference imaging is typically more computationally expensive but does well at removing extended non-variable objects such as galaxies, can detect variability in non-stellar profiles, and can perform better in crowded fields, while catalog searches require more accurate calibration but do not require template images and can be used to search for variability in data from different telescopes. For further discussion of difference imaging techniques, see [1], [7]. For a few examples of catalog searches, see [25], [8], [5] and [27]. A direct comparison of a catalog search for variability versus difference imaging for OGLE-II can be found in [30]. Generally, good results can be obtained as long as there is a match between the method's strengths and the science goals.

How does this all work in reality? There are many examples of synoptic surveys over the past few decades; a few are mentioned here. Due to hardware limitations (smaller telescopes, smaller fields of view), most past surveys have focused on one primary science goal. Still, the data from many of these projects has been adopted to search for many other kinds of variable or transient objects. Three surveys which were originally designed to search for potentially hazardous NEOs - the Lowell Observatory Near Earth Object Survey (LONEOS)[6], the Near-Earth Asteroid Tracking (NEAT) survey [22] and the Catalina Sky Survey (CSS)[17] - have (respectively) been used to discover RR Lyrae [20], joined in collaboration with the Palomar-Quest survey [10] to provide supernova candidates for the Nearby Supernova Factory [2], and serve as the source of data searched for all kinds of variability in the Catalina Real-Time Transient Survey (CRTS) [11].

Most of the observations of the Sloan Digital Sky Survey (SDSS) Stripe 82 were obtained primarily for a supernovae search [12], but Stripe 82 has provided a wealth of information about transients and variables of many kinds - see [5] and references therein. OGLE was designed to find microlensing events [28], but it and its successors OGLE-II and OGLE-III have found a similar multi-purpose success [29] from Cepheids [26] to transiting planets [4]. Other surveys originally designed to search for non-variable phenomena such as weak lensing at very faint limiting magnitudes have also been used to search for variable or transient objects - the Deep Lens Survey (DLS) [3] and the Subaru/XMM-Newton Deep Field (SXDF) [21] have both been searched for a wide range of variable and transient objects.

It is interesting to see the progression that has occurred, from searching for a particular target population (such as Type Ia SNe or Near Earth Objects or Trans-Neptunian Objects) to searching for a wide variety of objects. Current or imminent synoptic surveys, such as the Palomar Transient Factory [18], [23], Skymapper [16] and PanStarrs1 (PS) [9], and significantly larger future synoptic surveys, such as PanStarrs-4 and the Large Synoptic Survey Telescope (LSST) [15], [19], will be searching for all kinds of known transients and variables, as well as previously unknown objects where the requirements for understanding these 'unknown unknowns' are still developing.

Balancing the requirements for detecting and identifying a wide range of transients and variables is challenging, but these surveys are planning to meet the challenge with multi-color observations (typically some or all of u, g, r, i, z and y filters) over large portions of the sky (approximately 20,000 square degrees). Very large fields of view (7-10 square degrees) will allow these surveys to cover the sky on a rapid observational cadence; while the base plan is to cover all of the visible sky every few days (for LSST) or subsets of the visible sky every few days (PS4), it is likely that an entire range of revisit times between a few hours to several days may be explored for greater sensitivity to variability at a wider range of timescales. Data processing will be intensive - difference imaging will be used to detect transients and variables, and then software pipelines will run to identify moving objects, linking their detections into orbits. Accurate calibration of photometry and astrometry on a global reference scale requires more processing. LSST will send out public VOEvents describing the millions of transient and variable detections each night. Huge databases (upwards of 12 PB for LSST at the end of its 10 year survey) of objects and all of their detections will be created and available for queries. Users will be able to use this database to identify and classify the target samples of interest to them, as well as access the VOEvent stream to get real-time updates on targets.

These large synoptic surveys will open a whole new scientific window into the time-domain for astronomy, giving us the next evolutionary step and moving us fully into the 'action movie' of the universe.

References

[1] Alard, C. (2000). Image subtraction using a space-varying kernel. A&AS, 144:363–370. Alard, C. and Lupton, R. H. (1998). A Method for Optimal Image Subtraction. ApJ, 503:325–+.

[2] Aldering, G., et. al., (2002). Overview of the Nearby Supernova Factory. In J. A. Tyson & S. Wolff, editor, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, volume 4836 of Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, pages 61–72.

[3] Becker, A. C., Wittman, D. M., Boeshaar, P. C., Clocchiatti, A., Dell'Antonio, I. P., Frail, D. A., Halpern, J., Margoniner, V. E., Norman, D., Tyson, J. A., and Schommer, R. A. (2004). The Deep Lens Survey Transient Search. I. Short Timescale and Astrometric Variability. ApJ, 611:418–433.

[4] Bennett, D. P., et. al,. (2010). Masses and Orbital Constraints for the OGLE-2006-BLG-109Lb,c Jupiter/Saturn Analog Planetary System. ApJ, 713:837–855.

[5] Bhatti, W. A., Richmond, M. W., Ford, H. C., and Petro, L. D. (2010). Variable Point Sources in Sloan Digital Sky Survey Stripe 82. I. Project Description and Initial Catalog (0 hr $\leq\alpha\leq$ 4 hr). ApJS, 186:233–258.

[6] Bowell, E., Koehn, B. W., Howell, S. B., Hoffman, M., and Muinonen, K. (1995). The Lowell Observatory Near-Earth-Object Search: A Progress Report. In Bulletin of the American Astronomical Society, volume 27 of Bulletin of the American Astronomical Society, pages 1057–+.

[7] Bramich, D. M. (2008). A new algorithm for difference image analysis. MNRAS, 386:L77–L81.

[8] Bramich, D. M., Vidrih, S., Wyrzykowski, L., Munn, J. A., Lin, H., Evans, N. W., Smith, M. C., Belokurov, V., Gilmore, G., Zucker, D. B., Hewett, P. C., Watkins, L. L., Faria, D. C., Fellhauer, M., Miknaitis, G., Bizyaev, D., Ivezi □ c, Zˇ., Schneider, D. P., Snedden, S. A., Malanushenko, E., Malanushenko, V., and Pan, K. (2008). Light and motion in SDSS Stripe 82: the catalogues. MNRAS, 386:887–902.

[9] Chambers, K. C. and Pan-STARRS (2004). PS1 - The Prototype Pan-STARRS Telescope. In Bulletin of the American Astronomical Society, volume 36 of Bulletin of the American Astronomical Society, pages 1400–+.

[10] Djorgovski, S. G., Baltay, C., Mahabal, A., Rabinowitz, D., Drake, A., Donalek, C., Glikman, E., Graham, M., Williams, R., Ellman, N., Scalzo, R., Bauer, A., Nugent, P., and PQ Survey Team (2009). The Palomar-Quest Digital Synoptic Sky Survey: Summary and Initial Results. In American Astronomical Society Meeting Abstracts, volume 214 of American Astronomical Society Meeting Abstracts, pages 407.10–+.

[11] Drake, A. J., Djorgovski, S. G., Mahabal, A., Beshore, E., Larson, S., Graham, M. J., Williams, R., Christensen, E., Catelan, M., Boattini, A., Gibbs, A., Hill, R., and Kowalski, R. (2009). First Results from the Catalina Real-Time Transient Survey. ApJ, 696:870–884.

[12] Frieman, J. A., et. al, (2008) The Sloan Digital Sky Survey-II Supernova Survey: Technical Summary, Astron.J. 135:338-347

[13] Mendez, J., et. al (2008). The Sloan Digital Sky Survey-II Supernova Survey: Technical Summary. AJ, 135:338–347.

[14] Huber, M. E., Everett, M. E., and Howell, S. B. (2006). Color and Variability Characteristics of Point Sources in the Faint Sky Variability Survey. AJ, 132:633–649.

[15] Ivezic, Z., Tyson, J. A., Allsman, R., Andrew, J., Angel, R., and for the LSST Collaboration (2008). LSST: from Science Drivers to Reference Design and Anticipated Data Products. ArXiv e-prints.

[16] Keller, S. C., Schmidt, B. P., Bessell, M. S., Conroy, P. G., Francis, P., Granlund, A., Kowald, E., Oates, A. P., Martin-Jones, T., Preston, T., Tisserand, P., Vaccarella, A., and Waterson, M. F. (2007). The SkyMapper Telescope and The Southern Sky Survey. Publications of the Astronomical Society of Australia, 24:1–12.

[17] Larson, S., Beshore, E., Hill, R., Christensen, E., McLean, D., Kolar, S., McNaught, R., and Garradd, G. (2003). The CSS and SSS NEO surveys. In Bulletin of the American Astronomical Society, volume 35 of Bulletin of the American Astronomical Society, pages 982–+.

[18] Law, N. M., et. al, (2009). The Palomar Transient Factory: System Overview, Performance, and First Results. PASP, 121:1395–1408.

[19] LSST Science Collaborations, Abell, P. A., Allison, J., Anderson, S. F., Andrew, J. R., Angel, J. R. P., Armus, L., Arnett, D., Asztalos, S. J., Axelrod, T. S., and et al. (2009). LSST Science Book, Version 2.0. ArXiv e-prints.

[20] Miceli, A., Rest, A., Stubbs, C. W., Hawley, S. L., Cook, K. H., Magnier, E. A., Krisciunas, K., Bowell, E., and Koehn, B. (2008). Evidence for Distinct Components of the Galactic Stellar Halo from 838 RR Lyrae Stars Discovered in the LONEOS-I Survey. ApJ, 678:865–887.

[21] Morokuma, T., Doi, M., Yasuda, N., Akiyama, M., Sekiguchi, K., Furusawa, H., Ueda, Y., Totani, T., Oda, T., Nagao, T., Kashikawa, N., Murayama, T., Ouchi, M., Watson, M. G., Richmond, M. W., Lidman, C., Perlmutter, S., Spadafora, A. L., Aldering, G., Wang, L., Hook, I. M., and Knop, R. A. (2008). The Subaru/XMM-Newton Deep Survey (SXDS). V. Optically Faint Variable Object Survey. ApJ, 676:163–183.

[22] Pravdo, S. H., Rabinowitz, D. L., Helin, E. F., Lawrence, K. J., Bambery, R. J., Clark, C. C., Groom, S. L., Levin, S., Lorre, J., Shaklan, S. B., Kervin, P., Africano, J. A., Sydney, P., and Soohoo, V. (1999). The Near-Earth Asteroid Tracking (NEAT) Program: an Automated System for Telescope Control, Wide-Field Imaging, and Object Detection. AJ, 117:1616–1633.

[23] Rau, A., Kulkarni, S. R., Law, N. M., Bloom, J. S., Ciardi, D., Djorgovski, G. S., Fox, D. B., Gal-Yam, A., Grillmair, C. C., Kasliwal, M. M., Nugent, P. E., Ofek, E. O., Quimby, R. M., Reach, W. T., Shara, M., Bildsten, L., Cenko, S. B., Drake, A. J., Filippenko, A. V., Helfand, D. J., Helou, G., Howell, D. A., Poznanski, D., and

[24] Sullivan, M. (2009). Exploring the Optical Transient Sky with the Palomar Transient Factory. PASP, 121:1334–1351.

[25] Sesar, B., et. al. (2007). Exploring the Variable Sky with the Sloan Digital Sky Survey. AJ, 134:2236–2251.

[26] Soszynski, I., Poleski, R., Udalski, A., Szyman □ ski, M. K., Kubiak, M., Pietrzynski, G., Wyrzykowski, L., Szewczyk, O., and Ulaczyk, K. (2010). The Optical Gravitational Lensing Experiment. The OGLE-III Catalog of Variable Stars. VII. Classical Cepheids in the Small Magellanic Cloud. Acta Astronomica, 60:17–39.

[27] Telezhinsky, I., Eckert, D., Savchenko, V., Neronov, A., Produit, N., and Courvoisier, T. (2010). The catalog of variable sources detected by INTEGRAL I: Catalog and Techniques. ArXiv e-prints.

[28] Udalski, A., Szymanski, M., Kaluzny, J., Kubiak, M., and Mateo, M. (1992). The Optical Gravitational Lensing Experiment. Acta Astronomica, 42:253–284.

[29] Udalski, A., Szymanski, M. K., Soszynski, I., and Poleski, R. (2008). The Optical Gravitational Lensing Experiment. Final Reductions of the OGLE-III Data. Acta Astronomica, 58:69–87.

[30] Wozniak, P. (2008). Crowded Field Photometry and Difference Imaging. In Manchester Microlensing Conference.