

Everything we'd like to do with LSST data, but we don't know (yet) how

Željko Ivezić, Andrew J. Connolly and Mario Jurić

Department of Astronomy, University of Washington,
Box 351580, Seattle, WA 98195-1580, USA
email: ivezic@astro.washington.edu

Abstract. The Large Synoptic Survey Telescope (LSST), the next-generation optical imaging survey sited at Cerro Pachon in Chile, will provide an unprecedented database of astronomical measurements. The LSST design, with an 8.4m (6.7m effective) primary mirror, a 9.6 sq. deg. field of view, and a 3.2 Gigapixel camera, will allow about 10,000 sq. deg. of sky to be covered twice per night, every three to four nights on average, with typical 5-sigma depth for point sources of $r=24.5$ (AB). With over 800 observations in *ugrizy* bands over a 10-year period, these data will enable a deep stack reaching $r=27.5$ (about 5 magnitudes deeper than SDSS) and faint time-domain astronomy. The measured properties of newly discovered and known astrometric and photometric transients will be publicly reported within 60 sec after observation. The vast database of about 30 trillion observations of 40 billion objects will be mined for the unexpected and used for precision experiments in astrophysics. In addition to a brief introduction to LSST, we discuss a number of astro-statistical challenges that need to be overcome to extract maximum information and science results from LSST dataset.

Keywords. surveys, galaxies, stars: statistics

1. Introduction

The last decade has seen fascinating observational progress in optical imaging surveys. The SDSS dataset is currently being greatly extended by the ongoing surveys such as Pan-STARRS (Kaiser *et al.* 2010) and the Dark Energy Survey (Flaugher 2008). The Large Synoptic Survey Telescope (LSST) is the most ambitious survey currently planned in the visible band (for a brief overview, see Ivezić *et al.* 2008a). The unparalleled LSST survey power is due to its large étendue (see Figure 1).

The goals of the LSST are driven by four key science themes: probing dark energy and dark matter, taking an inventory of the Solar System, exploring the transient optical sky, and mapping the Milky Way. The LSST will be a large, wide-field ground-based system designed to obtain multiple images covering the sky visible from Cerro Pachón in Northern Chile. The system, with an 8.4m (6.7m effective) primary mirror, a 9.6 deg² field of view, and a 3.2 Gigapixel camera, will allow, on average, about 10,000 deg² of sky to be covered using pairs of 15-second exposures in two photometric bands every three nights, with a typical 5σ depth for point sources of $r \sim 24.5$. The system is designed to yield high image quality as well as superb astrometric and photometric accuracy †. The survey area will cover 30,000 deg² with $\delta < +34.5^\circ$, and will be imaged multiple times in six bands, *ugrizy*, covering the wavelength range 320–1050 nm. About 90% of the observing time will be devoted to a deep-wide-fast survey mode which will observe an 18,000 deg² region over 800 times (summed over all six bands) during the anticipated

† For detailed specifications, please see the LSST Overview Paper, Ivezić *et al.* (2008a), and the LSST Science Requirements Document (LSST Science Collaboration 2011)

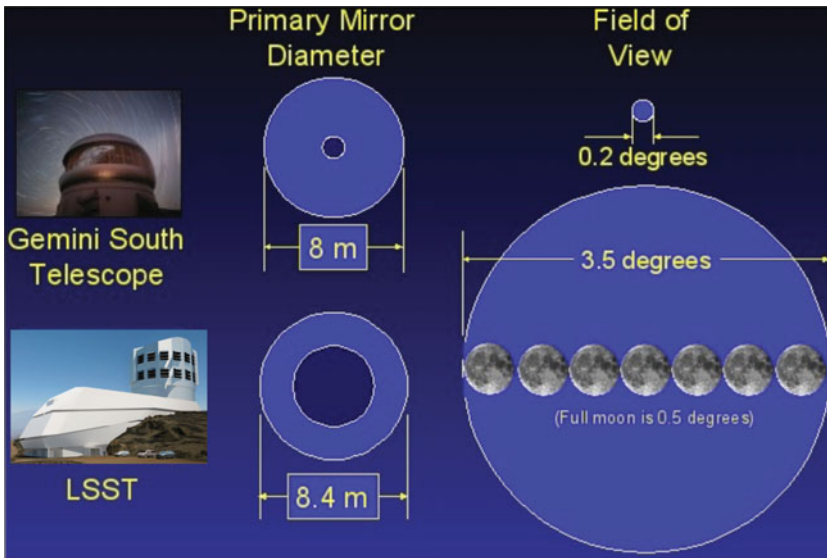


Figure 1. A comparison of the primary mirror size and the field-of-view size for LSST and Gemini South telescopes. The product of the primary mirror size and the field-of-view size, the so-called étendue (or grasp), a characteristic that determines the speed at which a system can survey a given sky area to a given flux limit, is much larger for LSST. Figure courtesy of Chuck Claver.

10 years of operations, and yield a coadded map to $r \sim 27.5$. These data will result in databases including about 20 billion galaxies and a similar number of stars, and will serve the majority of science programs. The remaining 10% of the observing time will be allocated to special programs such as a Very Deep and Fast time-domain survey, the details of which are still being defined. More details about various science programs that will be enabled by LSST data can be found in the LSST Science Book (LSST Science Collaboration 2009) and at the LSST website (www.lsst.org).

First light for LSST is expected at the end of 2019 with a small commissioning camera (144 Mpix), with the full 3.2 Gpix camera integrated in 2020. The construction phase of LSST, funded by the U.S. National Science Foundation and Department of Energy, started in 2015 and is progressing according to the planned schedule (see Figure 2).

2. LSST Data Analysis Challenges

The LSST project will deliver data products that will enable a large number of cutting-edge science programs (Jurić *et al.* 2016). Nevertheless, depending on the topic, the path from LSST data products to science results and journal papers may sometimes require additional challenging analysis work. These challenges, representative of the era of Big Data, stem from:

- Large data volumes (petabytes)
- Large numbers of objects (billions)
- Highly multi-dimensional spaces (thousands)
- Unknown statistical distributions
- Time-series data (irregular sampling)
- Heteroscedastic errors, truncated, censored and missing data
- Unreliable quantities (e.g. unknown systematics and random errors)



Figure 2. The inset in the top left corner shows a cut-away render of the LSST Observatory building. The rest of the figure shows a photograph of the LSST summit at the time of this Symposium (September 2016). First light for LSST is expected with a 144 Mpix engineering camera in 2019, and with the full 3.2 Gpix camera in 2020. For more photographs, see <https://www.lsst.org/gallery/image-gallery>

Everything we'd like to do with LSST data, but we don't know (yet) how is a catchy title but somewhat inaccurate. First, we most certainly do not include here “everything”, and second, we and our LSST colleagues already have at least some ideas for how to approach most of problems discussed below. We hope that this contribution will help motivate others to join us in this thinking, and to engage in the work needed to maximize LSST’s scientific yields.

To begin and stimulate this conversation, we have selected a few topics where substantial preparatory work is needed to optimally analyze datasets at the LSST scale. These are:

- (a) Interpreting spectral energy distributions (SEDs)
- (b) Identifying moving objects
- (c) Characterizing and classifying variable stars
- (d) Understanding systematic measurement uncertainties
- (e) Characterizing astrophysical simulations and astrophysical systematics
- (f) Devising new or enhanced algorithms to process LSST data.

We emphasize that many other members of LSST Science Collaborations contributed to the formulation of this, by all means, incomplete list. In the remainder of this section, we discuss these topics in a bit more detail.

2.1. Interpretation of spectral energy distributions (SEDs)

Efficient and robust interpretation of time-resolved multi-band photometry for “billions and billions” of objects is bound to yield unprecedented science results. A combination

of required measurement precision and relatively wide bandpasses will require careful interpretation of LSST data.

A broad-band photometric system, such as LSST, aims to deliver calibrated in-band flux

$$F_b = \int F_\nu(\lambda)\phi_b(\lambda)d\lambda, \quad (2.1)$$

where $F_\nu(\lambda)$ is specific flux of an object *at the top* of the atmosphere and $\phi_b(\lambda)$ is the normalized system response for the given band,

$$\phi_b(\lambda) = \frac{\lambda^{-1}S_b(\lambda)}{\int \lambda^{-1}S_b(\lambda)d\lambda} \quad (2.2)$$

(the λ^{-1} term reflects the fact that CCDs are photon-counting devices). Here, $S_b(\lambda)$ is the overall atmosphere + system throughput

$$S_b(\lambda) = S_b^{sys}(\lambda) \times S_b^{atm}(\lambda). \quad (2.3)$$

Numerous science programs can be cast as constraining the possible forms of the true SED $F_\nu(\lambda)$ given the measured broad-band fluxes, F_b , and the normalized system response, $\phi_b(\lambda)$, with $b = (u, g, r, i, z, y)$. Because of the integration over broad bandpasses, forward modeling using a trial SED (either empirical or model based) is typically superior to “correcting data” (fluxes, positions, sizes). Examples of such programs, where SEDs presumably depend on relevant astrophysical parameters, include

(a) photo- z algorithms: the observed galaxy and quasar SEDs depend on the redshift of an intrinsic SED (due to expansion of the universe, source evolution, and intergalactic extinction; see e.g. Bolzonella *et al.* 2000);

(b) photometric parallax for stars, where measured colors can be used to constrain the effective temperature and luminosity (e.g. Jurić *et al.* 2008);

(c) photometric metallicity for stars (trained using spectroscopic metallicities, see Ivezić *et al.* 2008b); and

(d) interstellar extinction along the line of sight for stars in the Milky Way disk (see, e.g., Berry *et al.* 2012).

There are a number of open issues that are being worked on by the community:

- What are the relative advantages and disadvantages of machine learning methods compared to methods based on fitting SED templates (both empirical and simulated). How can we incorporate ancillary data (and priors) within the photo- z methods, for example utilizing angular cross-correlation of photometric and spectroscopic samples of galaxies (Newman 2008)?

- What are the impacts of heteroscedastic noise, priors, and truncated and censored data?

- How much will “per-visit processing” of LSST data help (due to varying bandpasses $\phi_b(\lambda)$ because of the unavoidable variations in $S_b^{atm}(\lambda)$)?

- What is the best way to handle posterior probability density functions (pdfs), how much is gained compared to simple (e.g. maximum likelihood) point estimates, what are the optimal compression algorithms for pdfs, etc.?

- How should the parameter covariances be handled (the same question is also valid pretty much everywhere else below)?

2.2. Moving objects

The catalogs generated by LSST will increase the known number of small bodies in the Solar System by a factor of 10-100, among all populations (Jones *et al.* 2016). The

median number of observations for Main Belt asteroids will be on the order of 200-300, allowing sparse lightcurve inversion to determine rotation periods, spin axes, and shape information. The current strawman for the LSST survey strategy is to obtain two visits of the same field per night (each "visit" being a pair of back-to-back 15s exposures), separated by about 30 minutes, and covering the entire observable sky every 3-4 days throughout the observing season.

The main reason for two observations per night is to help association of observations of the same moving object from different nights, as follows. The typical distance between two nearby asteroids on the Ecliptic, at the faint fluxes probed by LSST, is a few arcminutes (counts are dominated by Main Belt asteroids). Typical asteroid motion during several days is much larger (of the order a degree or more) and thus, without additional information, detections of individual objects are "scrambled". However, with two detections per night, the motion vector can be estimated. The motion vector makes the linking problem much easier because positions from one night can be approximately extrapolated to future (or past) nights.

There are several interesting open questions regarding moving objects:

- Cadence optimization: are two visits per night really needed? Would perhaps a substantial increase in the computing power solve the association problem with just a single detection per night?
- How robust and efficient would be a full Bayesian approach for characterizing the orbits of asteroids (see, e.g., Virtanen *et al.* 2001)?
- How computationally hard is it to deploy shift-and-coadd method for KBOs and more distant objects on LSST-scale dataset?
- What are the most robust and efficient methods for sparse lightcurve inversion of several million asteroids?

2.3. Variable stars

Early in the survey, LSST will be discovering about 100,000 variable stars per night at high Galactic latitudes (Ridgway *et al.* 2014), and probably many more at low latitudes (but the forecast is less certain). The total number of variable stars to be discovered by LSST is of the order several hundred million (the total number of detected and measured stars will be about 20 billion). In addition, about 1000 new supernovae are expected to be discovered every observing night. A number of statistical questions need to be answered for the full exploitation of this dataset:

- How to best distinguish regular (periodic) from irregular variability when the data are sampled irregularly and when the variability may be wavelength dependent?
- How to distinguish short from long variability timescales?
- What are the best methods for the robust detection of variability, and for anomaly detection[†]? Recent developments in compressed sensing and deep learning have the potential to revolutionize the analysis of variability and transient detection. By exploiting the sparseness of the data and careful choice of the models that might be fit to these data, it maybe possible to characterize and classify sources in a way that is both flexible and robust to noise.
- What are the best methods for characterization and classification of a broad range of variability (especially in case of sparse data early in the survey)? How do machine learning methods compare to light curve template-based methods? Are there metrics

[†] Extensive tools for doing both template-matched and "model-independent" detection of variability have been recently developed in the context of LIGO.

that will enable a general classification scheme for identifying sources that might need follow up observations?

- Is it possible to further optimize the cadence to enhance discoveries and characterization of variable stars?
- What is the impact of heteroscedastic noise, astrophysical priors, and truncated and censored data?
- Can light curve and objects characterization and classification be done directly in database?

2.4. *Systematic measurement uncertainties*

Due to the large number of objects in LSST samples, many science programs, including cosmology, will be sensitive to systematic errors. In many cases the volume of the available data will mean that systematics are the dominant source of uncertainty (that is, given billions of objects measured a thousand times, how do we know that \sqrt{N} will still work in this regime?). The primary goals include:

- (a) ensuring that the astrometry can be measured with statistical and systematic errors at the milliarcsec level,
- (b) ensuring that the photometry can be measured with statistical and systematic errors at the millimag level,
- (c) measuring galaxy shapes (e.g., for use in cosmic shear analysis) with the PSF known across the focal plane to a level where the autocorrelation of PSF residuals is smaller than 10^{-7} .

These effects will need to be quantified as functions of position on the sky, position on the focal plane, observing conditions (e.g., atmospheric seeing, sky brightness), and object properties (e.g., brightness, colors, size). Some of the open questions include:

- How can the impact of unknown SEDs be quantified?
- What is the impact of the atmosphere (due to variable seeing and transmissivity, differential chromatic refraction and intrinsic stochasticity)?
- How can we robustly quantify both multiplicative and additive errors in galaxy shear measurements?
- How can we control systematic errors in photometric redshifts?

2.5. *Astrophysical simulations and astrophysical systematics*

The expected precision of the LSST measurements and their resulting constraints on cosmological and astrophysical models requires the development of simulation and modelling tools of equal or better precision. These tools will need to provide predictions for what the LSST will observe (in order to define effective survey strategies for the LSST), interpret the observations in the context of physically motivated models, and generate multiple realizations of a simulation to characterize the uncertainties or covariances within the cosmological and astrophysical constraints.

Simulations come in a number of different flavors: instrument simulations that model the optical performance of the LSST (e.g., the impact of the atmosphere, telescope, and camera on the image quality), cosmological simulations that model the evolution in structure within the universe, and mock or synthetic catalogs which provide realizations of the sky with appropriate distributions of the astrophysical properties of sources and their uncertainties.

The scale of the required simulations range from approximations of the technical parameters that describe the LSST (OpSim; Delgado *et al.* 2014), through single realizations of an image from an LSST sensor (DESC 2015) and large scale mock catalogs of asteroids

(Jones *et al.* 2016), to full simulations of representative volumes of LSST data (DESC 2015). Many of the required simulation tools are in development (e.g. Connolly *et al.* 2014). There remain, however, a number of challenges before these tools become widely accepted by the community.

- How do we support the generation of large scale simulations? The computational resources required to generate cosmological simulations, and in particular series of simulations for characterizing the covariance of cosmological models, are large and could exceed the resources available to individual investigators.

- How do we share simulations in a manner similar to the availability of observation data? Often the sizes of useful simulated data sets must exceed those of the observational datasets and are already approaching the PB scale. Transferring the generated simulations or mock catalogs from supercomputing centers to where they might be analyzed will stress academic network capacities.

- What is the impact of baryonic effects on dark matter halo profiles? The current generation of hydrodynamical simulations do not simulate large cosmological volumes. Approximations, where lower resolution or dark matter only models are used to identify regions of interest in the simulation that are then re-simulated at higher resolution, can lead to biases in any derived correlations as they are not representative volumes of the universe.

- What are the main feedback mechanisms in galaxy formation and what is the best way to handle nonlinear galaxy bias?

- How can we best address intrinsic alignments of galaxy shapes with the density field?
- How can we best extract the information about the evolution of the Milky Way galaxy using LSST measurements of 20 billion stars?

- How can we best extract the information about the evolution of the Solar System using LSST measurements of a few million asteroids?

2.6. LSST System Enhancements and New Algorithms

LSST is an automated *facility* that will deliver not only raw images, but also fully reduced data products (calibrated single-epoch images, multiple flavors of co-adds, and a variety of catalogs). Its cadence will be optimized to enable a balanced science return across the four key science themes (Section 1). Its data products have been designed to enable the derivation of a large fraction of those results without the need for end users to fully understand the details of the LSST instrument and data reduction, algorithms, or to begin from raw pixel data.

To make this possible, the LSST project is making a major investment in computing infrastructure, software, and algorithm development. Yet it is quite clear that more and better are always possible; even marginal improvement in performance (of both hardware and software) could yield significant additional science returns. While *some* of the open issues listed below are already being addressed by groups both within and outside the LSST construction project, substantial further research could be done.

Again, these are simply the most obvious examples; the list is by no means complete.

- Observing strategy (cadence) optimization can yield improvements in total open-shutter time for the survey, but also can improve the utility of angular and temporal sampling functions and dithering patterns (Delgado *et al.* 2014). It is, therefore, important to develop a scheduling algorithm that can efficiently address potential evolution of the LSST observing system and evolution of its science drivers. The LSST Project is

developing a scheduling algorithm that meets the survey requirements, but the complexity of the problem and the potential return on investment[†] argues for further research.

- LSST does not plan to deliver specialized crowded field reductions or catalogs; images of crowded regions of the Milky Way will be processed with the same code utilized elsewhere, though perhaps with different priors used in object detection and deblending stages (i.e., to a very good approximation, every object observed towards the Galactic center is a star). A purpose-built (multi-epoch capable) crowded field code capable of dealing with LSST source densities and data volumes would tremendously enhance the scientific return of LSST's Galactic dataset.

- No LSST data products have been explicitly designed to enable the detection and characterization of diffuse (e.g., ISM) or extremely low-surface brightness structures (e.g., LSB galaxies recently discovered by projects such as Dragonfly). Developing specialized codes to enable

such processing may add significant value to the LSST dataset.

- Complex galaxy models (e.g., tidal tails of merged galaxies) will not be fit by the standard LSST pipelines. Such a tool would greatly help in understanding gravitational potential around judiciously selected galaxies.

- Forward modeling of images on per visit basis (termed Multifit in LSST Data Management context) is superior to analysis of co-added images (because of varying observing conditions) and will be done by LSST. A particularly interesting problem is one of simultaneous forward modelling of data from *different* datasets (e.g., LSST and WFIRST). While there is substantial ongoing development (e.g. the Tractor code, see Lang *et al.* 2016), including within LSST Project, many statistical and other issues[‡] remain open and will require substantial further research to find the optimal approach.

- At the required precision level, the LSST point spread function (PSF) will depend on time, instrument state, source position, and source color (more precisely, on in-band SED shape); see Meyers & Burchat (2015). Robust and precise determination of the PSF will therefore be a rather non-trivial undertaking. The LSST project is required to characterize the PSF to the degree described in the LSST Science Requirements Document, but further improvements may be possible.

- A shift-and-stack algorithm (for co-adding images along arbitrary space-time trajectories), that could be efficiently deployed for large datasets would likely have a major impact on outer Solar System science. LSST Data Management (DM) system will not deliver shift-and-stack pipelines or data products, but these could be easily built on top of the open source code LSST DM will deliver.

- Image differencing will be used to detect transient sources in the LSST data stream. In order to control the false positive rate, new sophisticated algorithms will have to be developed to account for varying observing conditions (e.g., the treatment of differential chromatic refraction effects due to varying airmass, as well as color-dependent PSFs). At the same time, they will need to be fast enough to meet the requirement of delivering alerts within 60 seconds. The LSST project is developing these to enable its code deliverables, but broader research in this area would always be welcome.

- The SEDs and other properties of newly discovered transients will be poorly known initially. It is not clear yet what characterization and classification algorithms would be the best for separating the most interesting transients that require prompt followup from

[†] For example, just 1% effective improvement in LSST scheduling is roughly equivalent to ~\$4M in operational cost.

[‡] For example, blended objects present major algorithmic challenges and a discussion of their treatment, which is currently an open research area, would warrant a paper on its own.

the background of much more numerous transients which can be analyzed on much longer timescales without significant loss of science outcome.

- While there are well developed methods for the classification of light curves of variable stars (e.g. Richards *et al.* 2011), transient classification with sparse data is a much harder problem.

- Jointly processing data from LSST and other surveys (e.g., Euclid or WFIRST) would certainly result in a superior dataset than the one produced individually by either of these projects (for details see, e.g., Jain *et al.* 2015). It is not clear, however, how exactly to implement these ideas in practice, especially given that the survey overlap will be significantly truncated (either by position on the sky, e.g., for WFIRST, or by brightness, e.g. for Gaia).

- Finally, LSST data processing will be performed in the context of a relatively traditional HPC-like computing facility utilizing proven, low-risk, technologies (e.g., HT Condor, Pegasus). Similarly, LSST catalog data will be served to the users by way of relational databases (albeit of advanced, distributed, kind). Research into alternative models of processing (e.g., making use of the public cloud or workflow systems like Apache Spark) or data storage and serving (e.g., no-SQL databases, or next-generation experiments such as SciSQL) would be of great interest. If successful, these efforts could significantly enhance the ability of the community to perform affordable large-scale catalog computations or even image reprocessings.

A number of use cases above may be possible for the users to run at the LSST Data Access Centers. LSST has reserved approximately 10% of its total capacity to enable end-user analyses and generation of added value data products.

Furthermore, many of the use cases would be best tackled by enhancing existing LSST pipelines or building completely new functionality on top of the one already provided by LSST. All of the source code for the LSST pipelines will be publically available, enabling these kinds of endeavors.

Finally, LSST Operations have been built with the assumption that, in addition to the work within the facility, the community will make new discoveries and breakthroughs in areas of algorithms and data products. Such enhancements, developed by the community, can be incorporated into standard LSST processing, thus becoming a part of the official LSST alert streams and/or data releases.

3. Discussion

Due to the size and complexity of the LSST dataset, and the susceptibility of many of its major science programs to systematics in both measured quantities and astrophysical predictions, substantial preparatory work is required to enable the full exploitation of the LSST dataset. The bottleneck for science will not be the size of the dataset but instead our ability to extract useful and reliable information from the data.

Here we have summarized some of the most obvious research directions required to enhance the LSST science outcome. The main anticipated work areas include:

- advanced astronomical digital image processing,
- statistical modeling and analysis,
- data mining and machine learning,
- high performance computing,
- astrophysical simulations, and
- multi-dimensional and temporal data visualization.

LSST data analysis and the development of the fields of astro-informatics & astro-statistics will be closely intertwined. This synergy will open numerous opportunities for people with “Big Data” skills. Prospective LSST science users, across all disciplines, should collaborate and coordinate. By working jointly we can make the LSST great, and maximize the tremendous potential and science return of its dataset.

Acknowledgements

This material is based upon work supported in part by the National Science Foundation through Cooperative Agreement 1258333 managed by the Association of Universities for Research in Astronomy (AURA), and the Department of Energy under Contract No. DE-AC02-76SF00515 with the SLAC National Accelerator Laboratory. Additional LSST funding comes from private donations, grants to universities, and in-kind support from LSSTC Institutional Members. We thank Gregory Dubois-Felsmann for his careful reading and excellent comments.

References

- Berry, M., Ivezić, Ž., Sesar, B., *et al.* 2012, *Astrophysical Journal*, 757, 166
- Bolzonella, M., Miralles, J. M., & Pelló, R. 2000, *Astronomy & Astrophysics*, 363, 476
- Connolly, A. J., Angeli, G. Z., Chandrasekharan, S., *et al.* 2014, *Proceedings of the SPIE*, Volume 9150, id. 915014
- Delgado, F., Saha, A., Chandrasekharan, S., *et al.* 2014, *Proceedings of the SPIE*, Volume 9150, id. 915015
- DESC; Dark Energy Science Collaboration (DESC) Science Roadmap, 2015, http://lsst-desc.org/sites/default/files/DESC_SRM_V1.pdf
- Eyer, L., Evans, D. W., Mowlavi, N., *et al.* 2015, ArXiv:1502.03830
- Flaugher, B. 2008, In *A Decade of Dark Energy: Spring Symposium, Proceedings of the conferences held May 5-8, 2008 in Baltimore, Maryland. (USA)*. Ed. by N. Pirzkal & H. Ferguson.
- Ivezić, Ž., Tyson, J. A., Acosta, E., *et al.* 2008a, ArXiv:0805.2366
- Ivezić, Ž., Sesar, B., Jurić, M., *et al.* 2008b, *Astrophysical Journal*, 684, 287
- Jain, B., Spergel, D., Bean, R., *et al.* 2015, ArXiv:1501.07897
- Jones, R. L., Jurić, M., & Ivezić, Ž. 2016, *Proceedings of the IAU*, 318, 282, ArXiv:1511.03199
- Jurić, M., Ivezić, Ž., Brooks, A., *et al.* 2008, *Astrophysical Journal*, 673, 864
- Jurić, M., Kantor, J., Lim, K-T., *et al.* 2016, ASP Conf Ser. in press, ArXiv:1512.07914
- Kaiser, N., Burgett, W., Chambers, K., *et al.* 2010, *Proc. SPIE 7733, Ground-based and Airborne Telescopes III*, vol. 7733, 77330E
- Lang D., Hogg, D. W. & Mykytyn D. 2016, Astrophysics Source Code Library (ascl:1604.008)
- LSST Science Collaboration, 2009, LSST Science Book, arXiv:0912.0201
- LSST Science Collaboration, 2011, LSST Science Requirements Document, <http://ls.st/srd>
- Meyers, J. E. & Burchat, P. R. 2015, *Astrophysical Journal*, 807, 182
- Newman, J. A. 2008, *Astrophysical Journal*, 684, 88
- Reyes, R., Mandelbaum, R., Seljak, U., *et al.* 2010, *Nature*, 464, 256
- Richards, J. W., Starr, D. L., Butler, N. R., *et al.* 2011, *Astrophysical Journal*, 733, 10
- Ridgway, S. T., Matheson, T., Mighell, K. J., Olsen, K. A., & Howell, S. B. 2014, *Astrophysical Journal*, 796, 53
- Rockosi, C. M., Odenkirchen, M., Grebel, E. K., *et al.* 2002, *Astronomical Journal*, 124, 349
- Virtanen, J., Muinonen, K., & Bowell, E. 2001, *Icarus*, 154, 412
- Willman, B., Dalcanton, J. J., & Martinez-Delgado, D. 2005, *Astrophysical Journal*, 626, 85