



LARGE SYNOPTIC SURVEY TELESCOPE

# Large Synoptic Survey Telescope (LSST) LSST Data Quality Assurance Plan

Requirements for the LSST Data Quality Assessment Framework

**Tony Tyson, DQA Team, and Science Collaboration**

**LSE-63**

**Latest Revision: 2017-05-17**

**This LSST document has been approved as a Content-Controlled Document. Its contents are subject to configuration control and may not be changed, altered, or their provisions waived without prior approval. If this document is changed or superseded, the new document will retain the Handle designation shown above. The control is on the most recent digital document with this Handle in the LSST digital archive and not printed versions.**

## Abstract

LSST must supply trusted petascale data products. The mechanisms by which the LSST project achieve this unprecedented level of data quality will have spinoff to data-enabled science generally. This document specifies high-level requirements for a LSST Data Quality Assessment Framework, and defines the four levels of quality assessment (QA) tools. Because this process involves system-wide hardware and software, data QA must be defined at the System level. The scope of this document is limited to the description of the overall framework and the general requirements. It derives from the LSST Science Requirements Document [LPM-17]. A flow-down document will describe detailed implementation of the QA, including the algorithms. In most cases the monitoring strategy, the development path for these tools or the algorithms are known. Related documents are: LSST System Requirements [LSE-29], Optimal Deployment Parameters Document-11624, Observatory System Specifications [LSE-30], Configuration Management Plan [LPM-19], Project Quality Assurance Plan [LPM-55], Software Development Plan [LSE-16], Camera Quality implementation Plan [LCA-227], System Engineering Management Plan [LSE-17], and the Operations Plan [LPM-73].

## Change Record

Version	Date	Description	Owner name
1	2011-06-06	Initial version	Tony Tyson
2	2011-06-23	Added content	Tony Tyson
3	2011-07-08	General edits	Tony Tyson
4	2011-07-09	General edits	Tony Tyson
5	2011-07-10	General edits	Tony Tyson
6	2011-07-15	Added content	Tony Tyson
7	2011-07-31	Added content	Tony Tyson
	2017-04-13	Modernize style. Use BibTeX	Tim Jenness
	2017-04-16	Update content	Tony Tyson
7.1	2017-05-17	Implementation of LCR-943.	Tim Jenness



# Contents

<b>1 Introduction</b>	<b>1</b>
<b>2 LSST Data Quality Assessment Framework Structure</b>	<b>3</b>
<b>3 General Requirements for the LSST DQAF Design and Implementation</b>	<b>5</b>
<b>4 Specific Considerations for Each DQAF Level</b>	<b>7</b>
4.1 Level 0, software related . . . . .	7
4.2 Level 1, nightly data related . . . . .	7
4.3 Level 2, data release related . . . . .	8
4.4 Level 3, science related . . . . .	8
<b>5 Metrics of Metrics</b>	<b>9</b>
<b>6 Some Challenges and Opportunities</b>	<b>9</b>
6.1 Achieving acceptably low false transient alert rate . . . . .	10
6.2 Photometric redshift systematics . . . . .	11
6.3 Weak lens shear systematics . . . . .	12

# LSST Data Quality Assurance Plan

## 1 Introduction

With coverage of twenty billion objects, the LSST survey will enable breakthroughs due to the unprecedented precision of statistical studies, as well as the opening of the faint time domain. The vast amount of LSST data is a result of the required precision of the measurements and the required time-volume space coverage, driven by the novel science. Virtually all scientists who interact with the LSST data will make use of the photometric database (catalog) of objects and the measurements of their time behavior. Thus, these quantities must have a precision that is well characterized. In statistical investigations there is always a trade-off between completeness and precision. Different investigations usually favor different places along this relation. Indeed the broad range of LSST science recounted in the Science Book [11] spans the extremes. Some science relies more on completeness, some more on precision. Thus it will be necessary to accurately characterize this completeness-precision space for the data releases as well as in real time. Subtle sample selection effects as well as systematic errors will affect data quality. The current generation of surveys have often addressed systematic errors after the fact, during data analysis post data release. For LSST much of this understanding will emerge from explorations of the end-to-end simulations during R&D and construction. During operations the system will be designed to discover, report, and mitigate unanticipated errors and systematic effects. Understanding data quality issues to unprecedented levels prior to data release is for LSST an imperative.

Because errors in any part of the system can affect data quality, there is a need to obtain timely data quality information on the instrumental behavior of the Observatory, and to assist in the support of its maintenance. Since the LSST has only a single instrument, uptime concerns for the camera and the data system are paramount, and so it is important to support not only a basic level of detection of actual failures, but also to perform predictive maintenance, i.e., trend analysis on instrumental performance.

Finally there is the expected flood of transient object detections and alerts, requiring accompanying metadata capable of supporting robust (i.e. reproducible in tests) event classifications: this must be monitored and the tails of the distributions understood.

Several aspects of the LSST survey and resulting datasets drive the need for timely and extensive development of a comprehensive Data Quality Assessment Framework (DQAF):

- The quality of each image must be assessed promptly and thoroughly to track survey progress and the overall health of the LSST system.
- The large data volume and data rate prevents us from relying solely on deployment of traditional labor-intensive methods for assessing data quality.
- The distinctive design of the LSST as a single-instrument observatory with a dominant universal cadence observing model places a premium on the integrated uptime of the system. This requires the prompt detection and diagnosis, and - where possible - prediction of problems with the telescope or camera, facilitating timely maintenance.
- Many science programs have critical sensitivity to very small systematic errors. The needed system performance requires both appropriate image processing software, and the ability to discover subtle problems with the system (including hardware) or data taking strategy during the commissioning period, and during the course of the survey.
- Nearly real-time reporting of transient sources requires well defined triggering and their rapid and reliable characterization. A real-time data quality monitoring system is required to support these tasks, to minimize false transients (both false positives and false negatives) arising from flaws in the data.
- The unique deep-wide-fast features of the LSST survey will likely lead to discoveries of new rare phenomena. Their robust and efficient characterization will be intimately related to thorough and detailed understanding of the LSST data quality.

In summary, the LSST DQAF must be developed and validated prior to the commissioning period. DQAF needs to rely heavily on automated data analysis methods (such as data mining techniques for finding patterns in large datasets, and various machine learning regression techniques), and to be supported by modern data visualization tools. LSST is considered to be a lighthouse project for data-enabled science. It is likely that progress in developing tools and infrastructure for automated data quality assessment will have spin-off to e-Science generally. Already, exciting progress is being made on database technology that will have wide application.

The data processing software development process must include a full spectrum of QA/QC procedures, including design guidelines, code reviews, robust unit tests and coverage analysis, issue tracking, and system tests using a combination of simulated and - when available - real data from the Observatory. Software releases must be accepted through these procedures before installation in the production system. The LSST Software Development Plan

outlines some of the processes envisioned, and these will continue to be elaborated during the Final Design phase of the project.

Subsequent QA processes based on the acquired data must be integrated with the software development process, including the back-flow of tests based on lessons learned in the course of commissioning and the survey into the testing procedures in the development process. Some QA tools will be developed following a bottom-up design methodology. Indeed, many such tools are already deployed within the context of LSST Data Challenges. Nevertheless, the development process can be made more efficient and self-consistent by following general guidelines derived from the top-down approach as well. The main purpose of this document is to provide such guidelines.

## 2 LSST Data Quality Assessment Framework Structure

We describe here an overall organization of the framework for data quality assessment, from the science data stream point of view. Detailed requirements on the DQAF, including implementation details, will be developed in a separate document. The software engineering tools to track the software quality, as well as pipeline performance and job completion are discussed elsewhere.

The LSST DQAF will include four main components, which to some extent reflect the Level 1-3 structure of LSST data products. Level 0 QA is software development related, Level 1 QA relates to nightly operations, Level 2 QA relates to data releases, and Level 3 QA is science based.

- **Level 0 QA** includes the extensive and thorough testing of the DM subsystem during the pre-commissioning phase, as well as the tests of software improvements during the commissioning and operations phases (regression tests based on pipeline outputs and input truth). A common feature of Level 0 QA is the use of LSST simulations products, or any other dataset where the truth is sufficiently well known (e.g., the use of high-resolution observations from space telescopes to test star/galaxy separation algorithms). The main goal of Level 0 QA is to quantify the software performance against these known expected outputs (e.g., to measure the completeness and false positive rate for an object finder; to measure the impact of blended sources on pipeline outputs; to measure the performance of calibration pipelines and MOPS), and to test for algorithm implementation problems (a.k.a. “coding bugs”).



The whole system will be put to the test by processing the full set of one of the current big surveys (e.g., DES, HSC). Running one of these full data sets through, setting up the photometric system, and getting the alerts out will truly test the system. ComCam will play an important role as an end-to-end test within the LSST system. It is important to emphasize that the successful completion of Level 0 tests is a necessary but not sufficient condition for software acceptance. The Level 0 QA is a component of the comprehensive software QA/QC procedures envisioned in the LSST Software Development Plan, as mentioned above.

- **Level 1 QA** assesses the system status and data quality in real time during commissioning and operations. Its main difference compared to other observatory, telescope, and camera status reporting tools will be heavy reliance on the massive science imaging data stream (in addition to various telemetry and metadata generated by the subsystems). This level is tasked with nightly reporting of the overall data quality, including the nightly data products (difference images and transient source event stream) and calibration products. Real-time information about observing conditions, such as sky brightness, transparency, seeing, and the system performance, such as the achieved faint limit, will be delivered by Level 1 QA. Because the actual science data stream will be analyzed, Level 1 QA tools will be in a better position to discover and characterize subtle deterioration in system performance that might not be easily caught by tools employed by the telescope and the camera subsystems for self-reporting purposes. Level 1 QA also contributes to the Observatory's hardware performance monitoring and predictive maintenance capabilities.
- **Level 2 QA** assesses the quality of data products scheduled for the Data Releases, and provides quantitative details about data quality for each release (including the co-added image data products, and the properties of astrometrically and photometrically variable objects). This level also performs quality assessment for astrometric and photometric calibration, as well as for derived products, such as photometric redshifts for galaxies and various photometric estimators for stars. Subtle problems with the image processing pipelines and systematic problems with the instrument will be discovered with Level 2 QA.
- **Level 3 QA** quality assessment will be based on science analysis performed by the LSST Science Collaborations and other interested parties. Common features for tools at this level are sensitivity to subtle systematic issues not recognized by Level 2 QA, and feedback about data quality to the project by external teams. It is envisioned that especially useful Level 3 tools would be migrated to Level 2. Level 0-2 visualization and data ex-



ploration tools will be made available to the community. Examples of Level 3 analysis include testing of photometric redshifts with new spectroscopic samples unavailable to the project at the time of Data Release, comparison of photometric calibration with data from other large-area deep optical surveys, massive statistical analysis of systematic errors in image shape measurements, comparison of trigonometric and parallax measurements to catalogs from the Gaia mission, and the use of follow-up observations of transients to assess the quality of the LSST real-time classification engine.

### 3 General Requirements for the LSST DQAF Design and Implementation

Driven by the science requirements, a number of needed design and performance features are common to all four DQAF levels:

- Ability to incorporate auxiliary metadata when analyzing the science data stream.
- Ability to interpret information from the injection of artificial signals.
- Ability to automatically highlight problems and thus distinguish substandard data from the rest of data.
- Ability to track the consequences of data quality problems through data processing provenance - e.g., the ability to identify the source detections, object measurements, etc., that may be affected by a detected problem in the image data.
- Ability to rapidly and hierarchically explore possible causes of a problem after it has been identified and reported.
- Highly automated analysis, with a minimal reliance on expert human monitoring.
- Ability to monitor and statistically compare DQAF outputs across temporal, spatial (e.g. across the focal plane), and other axes.
- Ability to store, export and compare the results of DQAF analysis, including a user-friendly interface.
- Ability to import feedback from users, including problem discovery and fixes.
- Portability across the subsystems (e.g. the same image quality assessment tool should be available project-wide and should be deployable in different contexts).

- Ability to operate at different points in the overall data flow (e.g. an object count vs. size analysis could be performed using either outputs from the Data Release database, or using intermediate outputs from image processing pipelines, and then comparing the results).
- Flexibility to incorporate new tools or diagnostics, or to replace existing tools with improved versions (for example, new visualization and data mining tools).
- Availability of all tools employed in a particular DQAF level to all other DQAF levels (with possible exceptions applying to Level 3 QA). In particular, it is highly desirable that the same QA tests be available in the production context and in the software development context (i.e., in Level 0 QA).
- Documentation about assumptions made in various tests, detailed description of what they test for, as well as readily accessible information about all the tools at the user's disposal, should be an integral part of DQAF design.

Particular characteristics of LSST data will be multiple observations of the same sources, and a large number of objects in the sample. These properties enable numerous statistical methods to be applied for QA purposes. For example, very subtle systematic effects can be discovered by appropriately binning data along “interesting axes” (e.g., a photometric bias as small as a few millimag, or anomalous photometric noise, produced by a particular CCD, or amplifier, can be discovered by binning the differences between individual measurements and their average values by CCD coordinates). The ability for users to be able to specify the functional forms of those “interesting axes” would be very useful. User tools will be developed and available for creating additional functional forms that describe or produce the “interesting” quantities. In general, the relevant quantities that could drive systematic errors come in three flavors:

- Observing conditions (airmass, hour angle, seeing, sky brightness, photometricity, wind conditions);
- Instrumental parameters and behavior (CCD coordinates, CCD clock voltages, camera rotator angle, telescope elevation, system temperature);
- Object properties (brightness, colors, extendedness, shape, sky position in ecliptic and galactic coordinates).

The DQAF must support such statistical analyses on all the spatial and temporal scales of the survey. The issue of human – DQA pipeline interface design will be investigated and special attention will be paid to identifying problems with known causes vs those with causes that the software may not have been programmed to characterize.

## 4 Specific Considerations for Each DQAF Level

### 4.1 Level 0, software related

Level 0 tests are the first line of defense against faulty software. These tests and diagnostics will play a useful role in hardware validation as well. They will also serve as testbeds for detailed measurements of the software performance (completeness and false positive rate for object detection, biases in measured properties, the error distributions for measured quantities, etc). The same test suite can also be used to gauge the performance of newly developed algorithms. It is highly desirable to define a set of simulations that will serve as a “standardized test suite”, along with a standard chain of processing and analysis that is run automatically and answers a specific set of quantitative questions about software performance, in the manner of a regression test. These tests will be run as part of software release validation, both for Alert Production releases and for Data Release Production, and will also be available for use in the course of the software development process. A robust set of tests based on simulations will be essential during construction and early commissioning, but then must be supplemented with samples of real data from the Observatory as it becomes available.

### 4.2 Level 1, nightly data related

The Level 1 diagnostics will be one of the main sources of information for monitoring the overall state of the system (hardware and software) on a nightly basis. The camera diagnostics also operate on the actual data stream. However, because these Level 1 diagnostics will utilize the full results of a scientific data analysis, they will be in a better position than various telescope and camera status reporting systems to make decisions about the ultimate data quality. It is of paramount importance to have Level 1 tools completed and thoroughly tested well in advance of the commissioning period. These tools should be available earlier because they would be useful during the earliest system integration activities, and even during Camera subsystem integration. The survey cannot begin in earnest without these tools because they are needed for monitoring the quality of the real-time transient event stream and to map the survey progress.

### 4.3 Level 2, data release related

The Level 2 diagnostics will be among the main sources of information for establishing whether a Data Release is ready, and for quantifying its characteristics. Their outputs will also have a major impact on the understanding of systematic uncertainties in, and the interpretation of, scientific results obtained from the LSST data, and as such must be accompanied by appropriate publicly-available documentation. The Level 2 tools must also be able to incorporate auxiliary metadata, such as information from injection of artificial signals. Data exploration and visualization tools will have an important role.

Tools that are developed for Level 1 and 2 QA pipelines have to be made available to users as a support for Level 3 Data Product development and the associated DQA tools. This includes tools needed to efficiently explore multi-dimensional data.

### 4.4 Level 3, science related

It is likely that science teams would perform most of their Level 3 QA work using Data Release database access. The project should strongly encourage these teams to make their testing effort consistent with DQAF Levels 0-2. It will be desirable, over time, to incorporate at least some of the external tests and tools developed for Level 3 into the project-provided levels of the DQAF. For example, a complex statistical analysis that relies heavily on expert domain knowledge, and provides an excellent test for systematic errors in some measured quantity delivered by LSST, would be a good candidate for migration from Level 3 to Level 2 QA. An efficient way to convince external science teams to adopt, use, and contribute to the project-level QA effort is to provide timely and user-friendly documentation about the Level 1 and 2 tools. Data exploration tools should be made available to the community on an open-source basis.

LSST will produce large volumes of science data, largely precluding discovery of unanticipated anomalies by humans. The Data Management System produces derived products for scientific use both during observing (i.e. alerts and supporting image and source data) and in daily and periodic reprocessing. The periodic reprocessing also results in released science products. Analysis of the nightly data will also provide insight into the health of the telescope/camera system. However, by their very nature unanticipated anomalies cannot be efficiently discovered; one cannot write code to filter for something that is unanticipated. Yet it is just such “unknown unknowns” that can form either a major scientific discovery or (if they correspond to errors in the system) the misinterpretation of the science data. An automated data qual-

ity assessment system must include efficient searches for outliers in image data and unusual correlations of attributes in the database. This will involve aspects of data mining. Algorithms already exist for this kind of automated data quality assessment: spatio-temporal anomaly detection, and hyperspace cluster detection. Hints of a problem often come from anomalies in distributions. Variations around the mean values of measured parameters as a function of altitude and azimuth are an example. Finally, managing and testing new software builds will depend on efficient transfer of data between the production and development pipeline. The development systems will be tested on actual data from the production or operational pipeline. It is essential to ensure the development and operations teams have adequate and frequent communication.

## 5 Metrics of Metrics

How will we know if the DQA system is working? This has its own associated meta-metric. While individual metrics and associated algorithms must be designed to monitor each system quantity of interest, the overall metric we discuss here describes a method by which we can gauge the performance of the lower level metrics. Data quality requirements are not useful unless associated with a method and production system that will be used to test whether or not they are fulfilled in a given data set. Each level of DQA must be accompanied by an additional defined set of metrics which will assure that the DQA tasks are themselves being carried out effectively. In many cases these metrics will involve automated tests running in the background. For example, one method of validation is to inject anomalies into the data and measure the DQA recovery efficiency.

A test of this meta-metric is whether the LSST is meeting its overall scientific objectives. This is assessed at regular intervals (Operations Plan LPM-73), and the LSST Science Collaborations are preparing to push the envelope early in commissioning and operations. Indeed the most challenging goals discussed in the Science Book require unprecedented control of systematics and/or the development of novel informatics and data search algorithms. It is in these areas, briefly reviewed below, that DQA will be most critical.

## 6 Some Challenges and Opportunities

Below we list three challenging areas for LSST DQA which were identified by the LSST project and by the Astro-2010 Decadal Survey review process. A common feature of these three

examples is that they combine measurements from different filters for a large number of objects and across a large sky area, and thus enable discovery of subtle systematics that may go unnoticed when considering individual objects. Each of these is the subject of current R&D on the Project and in the LSST Science Collaborations, and it is expected that robust algorithms will be implemented during construction. The job for automated DQA will be to monitor the performance of these algorithms by subjecting them to frequent tests. These tests may take the form of injected events or objects in the LSST imaging data, or comparison of LSST object properties in the LSST database with external deeper data in selected areas (such as HST imaging).

## 6.1 Achieving acceptably low false transient alert rate

The science mission places stringent demands on the LSST's ability to rapidly and accurately detect, characterize, and classify varying and transient objects and to achieve a low false alarm rate. Here characterization means supplying object associated data for a source (photometric history, color, morphology, motion) and metadata (system status, etc) which could be used in classification of the object. Given the very high data volume produced by the LSST, the corresponding large number of detectable sources in each image (up to one million objects per visit), as well as the likelihood of entirely new classes of transients, the LSST will not be able to rely on traditional labor-intensive validation of detections, classifications, and alerts. It is estimated that about ten million transient alerts, including moving objects and variable stars, will be issued per night of LSST operations. While automated tools will be developed for data quality assessment, it is likely that visual checks of highly unusual events [alerted by the automated DQA system] and spot checks of routine events will be required to enhance the reliability of the alerts. These explorations will benefit from user friendly data exploration software which enables rapid correlation of user defined parameters in a selected sub-set of the high dimensional data, across multiple dimensions. Routine checks that the alert pipeline is working normally and generating alerts within specifications will also be required.

Characterization and preliminary object classification will be relatively straight forward for many sources, depending on how much time and wavelength coverage exists at the time of detection. The SRD requires tools enabling some level of classification as part of Alert Production: "The users will have an option of a query-like pre-filtering of this data stream in order to select likely candidates for specific transient type. Several pre-defined filters optimized for traditionally popular transients, such as supernovae and microlensed sources, will also be available." Of course it is likely that scientific discoveries will exist in the tails of these dis-

tributions as well as in unexpected parts of this multi-dimensional space. LSST will enlarge time-volume space by roughly a factor of 1000 over existing surveys, leading to the exciting prospect of discovery of new classes of objects. We must be ready for this discovery of the unexpected, and enable an efficient process of classification of transient sources. Of course some of these “unknown unknowns” may turn out to be errors in the LSST system, the detection and characterization of which is an imperative. Pursuing this via Levels 0-3 will help assure a low false alert rate. By “false” we mean either false positives (for example from pieces of diffraction spikes or *HII* regions in spiral arms) or false negatives. Automated monitoring of relevant system and data stream parameters, including efficiency of recovery of artificial events, will form the first line of defense against false alerts. Examples from current time domain surveys are the effects of known noisy areas of the focal plane, or detection of anomalous distributions in system metadata.

To achieve the levels of accuracy required, new algorithms for detection and classification must be created, as well as innovative automated techniques for alert filtering and validation. While not currently planned as part of the LSST data releases, the LSST project together with the science collaborations must push development of classification during R&D to “proof of principle.” As an example, simple machine learning and dimensional reduction algorithms are being developed by the Palomar Transient Factory collaboration [4]. These achieve efficiency through dimensional reduction. At greater than 96% classification efficiency, their samples achieve 90% purity by relying primarily on context-based features. Validated algorithms for automated detection of new types of transients should be part of level 2 DQA. This will form a second line of defense against false alerts. During operations the statistics of object classifications (and reclassifications) will be an important data quality diagnostic. The Zwicky Transient Factory will serve as a resource for test and development of these algorithms.

## 6.2 Photometric redshift systematics

One of the challenges in next generation sky surveys such as LSST is solving inverse problems using multi-dimensional petabyte databases. A key example of dimensional reduction is photometric redshifts: estimating redshifts of billions of galaxies based on many dimensional information (colors, brightness, sizes, and shapes). Multi-wavelength imaging photometry can be used to estimate the redshifts of every galaxy. This enables moderately narrow redshift intervals to be isolated so that distances (via the Hubble expansion) and the growth of structure can be charted as a function of cosmic time. The use of photometric redshifts does, however, come with an associated challenge, one that is common in all inversion problems:



the data are both noisy and incomplete (i.e., we do not have access to the full spectral energy distributions of all galaxies within a data set). The physics of photometric redshift determination implies that the measurement error distribution for redshifts has very long non-Gaussian tails. So called “catastrophic photo-z errors” are primarily caused by inappropriate galaxy type assignments resulting in the use of the wrong spectral energy distribution (SED). An issue of particular relevance to LSST is that little is known about the evolution of SEDs at higher redshifts for some types of galaxies.

Systematic errors in photometric redshift propagate to errors in cosmological parameters [1]. The principal components for understanding the impact of photometric redshifts are: estimating and minimizing systematics through the use of priors, calibration of the photometric redshift relation using spectroscopic training sets and by angular cross correlation between brighter spectroscopic and fainter photometric survey data, and characterizing the statistical uncertainties due to spatial and temporal variations in the survey progress. Photometric zero point errors [10] which vary across the sky can propagate to systematic errors in some cosmological parameters. Because LSST will cover different parts of the sky at different times, continuous automated DQA will be implemented for these metrics. An additional challenge for photometric redshifts is blending of barely resolved multiple galaxies; thus the LSST deblender algorithm must have reasonable efficiency for tagging blends at the  $i = 25.3$  mag limit of the LSST main survey “Gold Sample” of galaxies. Systematic errors in galaxy color morphology vs those of stars occur at high airmass, and algorithms which leverage these spatio-color correlations will be helpful. As mentioned above, monitoring derived parameters as a function of other variables, such as position in the sky or system parameters, is an important activity for Level 2 and 3 QA.

### 6.3 Weak lens shear systematics

Mapping dark matter and probing dark energy [two prime science drivers for LSST] make use of the weak gravitational lens shear of background galaxies by foreground mass structures. Measurements of cosmic shear correlate the shapes (shears) of pairs of galaxies separated on the sky. Cross correlations between galaxy shears in different redshift samples (cosmic shear tomography) as a function of angular separation on the sky is a sensitive diagnostic of cosmology when combined with correlations of the galaxy locations (baryon acoustic oscillations). The hemisphere sky coverage of LSST is needed in order to achieve the required statistical precision in these shear correlations, and to suppress cosmic variance. For the LSST “gold” sample of 4 billion galaxies, the resulting random component of the shear cross cor-

relation noise level is about  $3 \times 10^{-7}$  over an angular range up to several degrees. It is thus important that the systematic component be less than about 30% of this noise. The requirement then is that the galaxy shear extraction algorithm (and system hardware) be capable of delivering this level of galaxy shear systematics residual. Because of the stochastic nature of galaxy shape shot noise, the shear errors in a large sample are dominated by PSF errors when extrapolated to the galaxy positions (together with errors in model fitting, given the PSF). Stars whose core flux is brighter than the galaxies in the weak lens sample are used for PSF measurement. However, thick fully-depleted CCD exhibit a “brighter-fatter” PSF effect, changing pixel shapes, locations, and sizes on and near these PSF calibration stars. This could be removed by the pixel processing pipeline by implementing corrections based on CCD device physics models of sufficient accuracy to enable on-sky updates of parameters, or using a well-calibrated empirical model.

Faint galaxy shear is thus contaminated by PSF variations. The PSF must be precisely mapped for each exposure and for each CCD. The distribution of PSF residuals for up to 20,000 stars per pointing will be monitored, and is an important diagnostic of data quality. Consistency with the wavefront sensing solution will be another diagnostic. Several algorithms have been suggested to reduce PSF systematics in galaxy shear extraction using multiple exposures of the same field. The naive use of such data would be to construct a single “co-added” image with higher signal-to-noise, and then measure the shear correlation function by averaging over all pairs of galaxies. The LSST lens pipeline will more likely analyze the full “data cube” by fitting, for each galaxy, a single model which best matches the  $n$  measurements of that galaxy in the survey, when convolved separately with the  $n$  corresponding PSFs (the *MultiFit* method, under development). A faster algorithm (called *StackFit*) co-adds the  $n$  weighted PSF eigenfunctions and fits a model to the co-added galaxy image [9].

Automated statistics of galaxy-by-galaxy consistency of the shear should provide powerful metrics of shear systematics. An example is automated calculation of a shear cross correlation called “B-mode” which in the absence of systematic PSF errors should be zero. We should expect to detect some sky tiling effects on weak lens shear systematics due to residuals from PSF corrections at the field edges of overlapping visits taken under different seeing conditions. Achieving the cosmology promise of LSST will require strict control of systematic errors on large scales. Repeatability of shear patterns on revisits and in overlapped dithered exposures will be a useful Level 3 metric. Special dithering during observing will help suppress PSF shear systematics, by rotating the camera in addition to x-y dithering and observing the same area at various HA such that a full 180 degree is covered over the full re-visit sample. This is discussed in the recent LSST Observing Strategy whitepaper.

## References

- [1] Abrahamse, A., Knox, L., Schmidt, S., et al., 2011, *ApJ*, 734, 36 (arXiv:1011.2239), doi:10.1088/0004-637X/734/1/36, ADS Link
- [2] **[LSE-16]**, Allsman, R., Dubois-Felsmann, G., Kantor, J., 2009, *LSST Software Development Plan*, LSE-16, URL <https://ls.st/LSE-16>
- [3] **[LPM-19]**, Angeli, G., McKercher, R., 2015, *Change Control Process*, LPM-19, URL <https://ls.st/LPM-19>
- [4] Bloom, J.S., Richards, J.W., Nugent, P.E., et al., 2012, *PASP*, 124, 1175 (arXiv:1106.5491), doi:10.1086/668468, ADS Link
- [5] **[LSE-17]**, Claver, C., Angeli, G., Selvy, B., 2016, *Systems Engineering Management Plan*, LSE-17, URL <https://ls.st/LSE-17>
- [6] **[LSE-29]**, Claver, C.F., The LSST Systems Engineering Integrated Project Team, 2017, *LSST System Requirements (LSR)*, LSE-29, URL <https://ls.st/LSE-29>
- [7] **[LSE-30]**, Claver, C.F., The LSST Systems Engineering Integrated Project Team, 2018, *Observatory System Specifications (OSS)*, LSE-30, URL <https://ls.st/LSE-30>
- [8] **[LPM-17]**, Ivezić, Ž., The LSST Science Collaboration, 2011, *LSST Science Requirements Document*, LPM-17, URL <https://ls.st/LPM-17>
- [9] Jee, M.J., Tyson, J.A., 2011, *PASP*, 123, 596 (arXiv:1011.1913), doi:10.1086/660137, ADS Link
- [10] Jones, R.L., Padmanabhan, N., Ivezić, Z., et al., 2010, In: *Observatory Operations: Strategies, Processes, and Systems III*, vol. 7737 of Proc. SPIE, 77371F, doi:10.1117/12.857743, ADS Link
- [11] LSST Science Collaboration, 2009, ArXiv e-prints (arXiv:0912.0201), ADS Link
- [12] **[Document-11624]**, LSST Science Council, 2011, *Optimization of LSST Deployment Parameters*, Document-11624, URL <https://ls.st/Document-11624>
- [13] **[LCA-227]**, Nordby, M., Kurita, N., O'Neill, F., Marsh, D., 2014, *LSST Camera Quality Implementation Plan*, LCA-227, URL <https://ls.st/LCA-227>
- [14] **[LPM-55]**, Sweeney, D., McKercher, R., 2013, *Project Quality Assurance Plan*, LPM-55, URL <https://ls.st/LPM-55>



[15] **[LPM-73]**, Wolff, S., 2013, *Operations Plan*, LPM-73, URL <https://ls.st/LPM-73>