Commissioning, and Pre-Operations Data Policy & Serving

Bob Blum, Chuck Claver, Leanne Guy
Outline

- Commissioning Overview
- Commissioning Community Inputs
- Commissioning Data Releases
- Early Operations Supporting Data Releases
- Engagement with the Project Commissioning Effort
The commissioning plan (LSE-79) is structured in phases

Phase 0: Pre-Commissioning (happening now)
- System verification test & procedure definitions
- Test analysis script development using the LSST code base
- Procedure & script validation with simulation tools
- Special hardware/software development and logistics
- Summit Facility Camera Clean Room Builds

Phase I: Early System AI&T with ComCam - 6 months
- initial interface testing
- telescope alignment and AOS control
- sub-scale DM pipeline testing

Phase II: Full System AI&T with Science Camera -
- LSSTCam shipping and re-verification
- LSSTCam-Telescope integration, alignment & AOS control
- full scale DM pipeline testing

Phase III: Science Validation
- mini-Survey 1 – image differencing templates and alert production
- mini-Survey 2 – 10-year image stack and data release processing
- Operations readiness

Critical sky positions where some form of “truth” is known have been identified.

Analysis tooling development with LSST Science Platform is underway
Planned On-sky Observing Campaigns

<table>
<thead>
<tr>
<th>Early Science Validation with ComCam</th>
<th>Early Science Validation with LSSTCam</th>
<th>Science Validation Surveys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Installation and Initial Testing</td>
<td>Installation and Initial Testing</td>
<td></td>
</tr>
<tr>
<td>Engineering focus, algorithm testing, instrument signature removal</td>
<td>Engineering focus, algorithm testing, instrument signature removal</td>
<td></td>
</tr>
<tr>
<td>3-4 months</td>
<td>3 weeks</td>
<td>3 weeks</td>
</tr>
<tr>
<td>Key Performance Metrics</td>
<td>Key Performance Metrics</td>
<td>Survey 1: Wide Area</td>
</tr>
<tr>
<td>Image quality, depth, astrometry, photometry</td>
<td>Image quality, depth, astrometry, photometry</td>
<td>Template generation</td>
</tr>
<tr>
<td>4 weeks</td>
<td>Survey 2: Full Depth</td>
<td>Survey 2: Full Depth</td>
</tr>
<tr>
<td>20-year Depth Test</td>
<td>10-year survey depth in reference fields overlapping with deep external imaging and spectroscopy datasets</td>
<td></td>
</tr>
<tr>
<td>Exploring range of conditions</td>
<td>6 weeks</td>
<td>Survey 1: Wide Area</td>
</tr>
<tr>
<td>4 weeks</td>
<td>Scheduler Tests</td>
<td>Real-time alert production</td>
</tr>
<tr>
<td>Nominal cadence, ToOs, environmental conditions</td>
<td></td>
<td>3 weeks</td>
</tr>
<tr>
<td>4 weeks</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Science Validation Surveys

- Two 6-week continuous scheduler-driven surveys exercising the prompt and data release processing science pipelines
- Comprehensive characterization of bulk data acquired under nominal observing conditions
- Identifying corner cases with the aid of a larger statistical sample

<table>
<thead>
<tr>
<th></th>
<th>Early Science Validation with ComCam</th>
<th>Early Science Validation with LSSTCam</th>
<th>Science Validation Surveys</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-4 months</td>
<td>Installation and Initial Testing</td>
<td>Installation and Initial Testing</td>
<td>Survey 1: Wide Area</td>
</tr>
<tr>
<td></td>
<td>Engineering focus, algorithm testing, instrument signature removal</td>
<td>Engineering focus, algorithm testing, instrument signature removal</td>
<td>Template generation</td>
</tr>
<tr>
<td>4 weeks</td>
<td>Key Performance Metrics</td>
<td>Key Performance Metrics</td>
<td>Survey 2: Full Depth</td>
</tr>
<tr>
<td></td>
<td>Image quality, depth, astrometry, photometry</td>
<td>Image quality, depth, astrometry, photometry</td>
<td>10-year survey depth in reference fields overlapping with deep external imaging and spectroscopy datasets</td>
</tr>
<tr>
<td>4 weeks</td>
<td>20-year Depth Test</td>
<td>20-year Depth Test</td>
<td>Survey 1: Wide Area</td>
</tr>
<tr>
<td></td>
<td>Exploring range of conditions</td>
<td>Exploring range of conditions</td>
<td>Real-time alert production</td>
</tr>
<tr>
<td>4 weeks</td>
<td>Scheduler Tests</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nominal cadence, ToOs, environmental conditions</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PST Briefing to Science Collaboration Chairs - March 27, 2018
Science Validation Survey 1: Wide Area

Objectives
- Validate template building with Data Release Processing pipeline
- Alert Processing, real-time alert generation
- Monitor survey progress over wide area to test observation simulations

Observations (location and distribution on the sky TBD)
- ~1600 deg\(^2\) x 15 visits x 6 filters x 2 phases (~30K visits, ~40 nights)
- Phase 1: observations for template generation (3 weeks)
- Phase 2: observations of same area for alert production (3 weeks)
- Phases separated by 6 weeks to allow for astrophysical evolution and template processing (Science Validation Survey 2 scheduled between phases)

Additional Considerations
- Use dithered pointings to match Wide-Fast-Deep pattern
- Use large sky area to explore edge cases (bright stars, high source densities, etc.)
Science Validation Survey 2: 10-yr Depth

Objectives
- Focus on Data Release Products at full survey depth
- Data quality characterization beyond the SRD
- Template generation and real-time alert production (more rapid cadence may enable unique tests)

Observations (location and distribution on the sky TBD)
- ~300 deg² x 825 visits across 6 filters (~30K visits, ~40 nights)
- Select fields to overlap with external reference fields
- Scheduler used to optimize data quality across fields

Additional Considerations
- Use dithered pointings to match Wide-Fast-Deep pattern
- Option to select adjoining fields to form larger contiguous full-depth regions
- Alert Processing studies would benefit from early template generation
# Key Data Acquisition Dates: Current Forecast

<table>
<thead>
<tr>
<th>Milestone</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start of On-Sky Data from Auxiliary Telescope</td>
<td>Oct. 2019</td>
</tr>
<tr>
<td>Start of On-Sky &amp; Calibration Data with ComCam</td>
<td>Oct. 2020</td>
</tr>
<tr>
<td>Sustained Observing with ComCam</td>
<td>Feb. 2021</td>
</tr>
<tr>
<td>Start of On-Sky &amp; Calibration Data with LSSTCam</td>
<td>July 2021</td>
</tr>
<tr>
<td>Sustained Observing with LSSTCam</td>
<td>Oct. 2021</td>
</tr>
<tr>
<td>Start of Science Verification mini-Surveys</td>
<td>Dec. 2021</td>
</tr>
<tr>
<td>Operations Readiness Review</td>
<td>March 2022</td>
</tr>
</tbody>
</table>
How, where and what are we going to observe?

<table>
<thead>
<tr>
<th>Early Science Validation with ComCam</th>
<th>Early Science Validation with LSSTCam</th>
<th>Science Validation Surveys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Installation and Initial Testing</td>
<td>Installation and Initial Testing</td>
<td>Survey 1: Wide Area</td>
</tr>
<tr>
<td>Engineering focus, algorithm testing, instrument signature removal</td>
<td>Engineering focus, algorithm testing, instrument signature removal</td>
<td>Template generation</td>
</tr>
<tr>
<td>3-4 months</td>
<td>3 weeks</td>
<td>Survey 2: Full Depth</td>
</tr>
<tr>
<td>Key Performance Metrics</td>
<td>Key Performance Metrics</td>
<td>10-year survey depth in reference fields overlapping with deep external imaging and spectroscopy datasets</td>
</tr>
<tr>
<td>Image quality, depth, astrometry, photometry</td>
<td>Image quality, depth, astrometry, photometry</td>
<td>6 weeks</td>
</tr>
<tr>
<td>4 weeks</td>
<td>4 weeks</td>
<td>Survey 1: Wide Area</td>
</tr>
<tr>
<td>20-year Depth Test</td>
<td>20-year Depth Test</td>
<td>Real-time alert production</td>
</tr>
<tr>
<td>Exploring range of conditions</td>
<td>Exploring range of conditions</td>
<td>3 weeks</td>
</tr>
<tr>
<td>Scheduler Tests</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nominal cadence, ToOs, environmental conditions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 weeks</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Field Identification - on going but not done
**Input form for suggestions:**

[https://docs.google.com/forms/d/11FXz5noEwYPYtjqCziv7ZQhTlabFE-LFAQHVINU4jWI/edit](https://docs.google.com/forms/d/11FXz5noEwYPYtjqCziv7ZQhTlabFE-LFAQHVINU4jWI/edit)

<table>
<thead>
<tr>
<th>QUESTIONS</th>
<th>RESPONSES</th>
</tr>
</thead>
</table>

**LSST Commissioning Input**

form to propose commissioning observations

**Email address** *

Valid email address

This form is collecting email addresses. [Change settings](https://help.google.com/forms/answer/9232924?hl=en)

**Science Collaboration** *

Leading Science Collaboration (if multiple choose other and specify in text box)
All data release are seen as operations or pre operations activities - the MREFC project, including commissioning, has no requirement to serve data to the community.
Commissioning Data for Community

- The LSST Commissioning Team will generate data products from the observations taken in commissioning. The Commissioning team will verify, validate, and characterize these products, in order to test for the operations readiness of the system.

- The LSST Operations team plans to use the commissioning data products to develop and test its procedures for releasing, and then supporting the use of, LSST data.

- Preparing for each data release “scenario” in commissioning and science validation will take time and effort by the ops team.
- At a minimum LSST Operations envisions serving flat FITS files of images and catalog data from early commissioning.

- Operations expects that the data products generated from the commissioning and science validation (SV) observations carried out with both ComCam and LSSTCam will be made available through the (perhaps limited functionality during commissioning) LSST Science User Interface around 6-9 months after the observations are completed. Earlier previews of the data may also be possible. The final SV survey data products will be served to the community via the full LSST Science Portal (which includes the SUI).

- The schedule, and scope, of these scenarios is likely to evolve as commissioning proceeds: the Ops team is preparing a data release scenarios document (LSO-011), and will keep the science collaborations informed!
Commissioning Data for Community

- Documentation generated (drafted?) by commissioning team. Operations will (improve? And) serve documents and interface with community to support commissioning/verification on best effort basis

- All commissioning data products and services evolve toward full survey quality and content as we approach October 1, 2022.
Data Release Scenarios

- Scenario is a possible data release schedule or timeline
- Current plans include 3 scenarios for pre-operations (ComCam, LSSTCam, SV mini-Surveys)
- Release may start at flat files, will evolve to fully supported Data Access Center with LSP (LDM-554)
- Plan to release data through LSP ASAP to gain operational experience and user feedback
- Scenario 1, first image data released with 6 months following ComCam data complete. Catalogs 3 months later
- Scenario 2, 3 full data preview released 6 months after data taking complete. (See S18)
Data Quality

- Responsibility of commissioning team
- Which data are released will be subject to Project review - e.g. not all commissioning data will be released
- Data will be released as shared risk for scientific use
- Feedback on data quality/issues to commissioning team will go through some TBD channel
Scenarios

- ComCam Q2 2022
  • Data taking ends June 2021
  • Start releasing data (images) Dec 2021
  • Full data release Mar 2022

- LSSTCam Q3 2022
  • Data taking ends Dec 2021
  • Full data release June 2022

- SV mini-Surveys Q4 2022
  • 1-3 surveys
  • Data taking ends Mar 2022
  • Deliver first MS six months after finished (Sep 2022)
Prompt Data Products

- Prompt Processing pipelines including, precovery, MOPS, and alert generation and distribution will be run in commissioning
- Difference images will be produced as templates become available to subtract from
- Expect to produce normal PPDB catalogs
  - < 12 months of DIASource records.
  - see DPDD for details
- If DRP data products are available in Qserv, could attempt linking - depends on DRP timeline
Alerts

- Early alert stream data will also be distributed, possibly with substantial latency, to familiarize the community with the planned format and content of the alert stream.

- A best-effort alert stream may be distributed around Aug 2021.
  - ComCam +3 months: Produce sample commissioning alerts
  - LSSTCam M/S +1 month: Produce sample commissioning alerts

- Distribution Mechanism: Kafka/AVRO based system
  - Currently being tested on ZTF (and ingested by ANTARES and Lasair).

- See [LDM-612](#) for more information on alerts and brokers.
Alert packet contents in Commissioning

Each Alert (a VOEvent packet) will contain at least the following:

- **alertID**: An ID uniquely identifying this alert. It can also be used to execute a query against the Level 1 database as it existed when this alert was issued.

- **Level 1 database ID**

- **Science Data**:
  - The DIASource record that triggered the alert, as well as the filterName and programId of the corresponding Visit
  - The entire DIAObject (or SSObject) record
  - Previous 12 months of DIASource records
  - Matching object IDs from the latest Data Release, if they exist, and 12 months of their DIASource records

- Cut-out of the difference image centered on the DIASource (10 bytes/pixel, FITS MEF)

- Cut-out of the template image centered on the DIASource (10 bytes/pixel, FITS MEF)

- Probably only for known asteroids
  - Will be < 12 months

- There will be no DR in commissioning, we will do what we can with the DRP data products
Moving Objects Pipeline

- Expect to run MOPs in the Mini Surveys Q4 2022
- Validate the the Solar System products pipeline
  - Test the linking software works
  - Test the interface to the Minor Planets Center
- Prompt Data Products -> SSOBJECT Table, see DPDD
  - Possibly only for known asteroids
- Should not expect operations-era latencies of 24h for solar system data products in commissioning
- Any discoveries can be released as a byproduct of this activity
- Will run improved versions of the LSST science pipelines that we run today for the HSC reprocessing
- Expect many improvements in algorithms throughout commissioning. Expect:
  - improvements in quality
    - but the form of outputs not expected to change
  - improvements in QA and V&V processes
- Catalogs in Qserv in Science Data Model format
  - See DPDD for table descriptions: ls.st/LSE-163
- All data release are seen as operations or pre operations activities - the MREFC project, including commissioning, has no requirement to serve data to the community.

- During both commissioning and operations the plan will have to be adjusted according to the commissioning schedule, success of the data reductions (quality of the data products), resources available for preparing and supporting a release, and the need for incremental releases (i.e. will a new intermediate release be a significant improvement on the previous release).

- The capability level of a Data Access Centre will also have to be considered - the final platform is well understood (LDM-554) but at what point specific features may become available is TBD in commissioning.

- We do not present a release schedule; instead we present a release scenario which may or may not play out depending on the many unknowns and caveats on the various data releases and available resources.
The Project welcomes external groups to add value and complement the existing LSST Commissioning Team’s capabilities.

The Project reserves the right to decline any group’s offer for any reason.

The Project can terminate a group’s affiliation at any time due poor performance or undue overhead.

The Project will not rely on the contributions of external groups to fulfill the core commissioning requirements to the Operations Readiness Review.
Engagement with LSST Commissioning

- I can’t pay you
- You would be signing up to do real work on a schedule as directed by the LSST Project Commissioning Team – terms set by MOU
- You are expected to spend time at one of the primary LSST operations sites prior to and during commissioning

What’s in it for you?
- Open access to everything we learn during commissioning
- Deep understanding of LSST data before the survey starts
Basic Guidelines for Engagement

1) All agreements to participate in LSST commissioning activities will be done though a formal signed MOU between the LSST Project and the group’s home institution.

2) No scientific publication on the commissioning data shall be made prior to the data going public. Technical publications will be allowed.

3) Each group must identify a senior person to act as the primary point of contact with the LSST commissioning team. This person will ensure that assigned work to their group is completed on the schedule required by the Project and ensure the highest of quality.

4) Participating individuals are expected to spend time at one of three primary LSST sites:
   1) Chile – either La Serena and/or Cerro Pachón;
   2) LSST offices in Tucson; and/or
   3) National Center for Super Computing Applications (NCSA) at Champaign-Urbana, IL.

5) Participation will be based on value added and capabilities to complement the existing LSST Commissioning team.
Conduct anomaly analysis of the Engineering Facility Database. This analysis package is meant to apply Machine Learning/Deep Learner and other AI or similar algorithms to search for otherwise undetected anomalies in the system performance and correlate these with properties of the image data and catalog parameters.

Extended analysis to characterize system performance at the margins of operational parameter space. Over the course of commissioning the LSST Telescope and Camera will be operated over a wide range of environmental conditions. This analysis package is meant to characterize the multiple performance metrics, of both scientific and technical kinds, and explore performance correlations with respect to environmental parameters.

Evaluate operational configurations of the observatory to determine optimum performance. The LSST Observatory has been design with many degrees of freedom built in to optimize system performance to respond to different observing conditions. This analysis package is meant to analyze where optimal operational parameters, configurations and procedures lie for the as built system.