Commissioning Science Verification and Validation

Keith Bechtol & Andrew Connolly
Science Verification Leads

NSF/DOE Joint Status Review
August 27-30, 2019

Outline

- Overview of Commissioning SV
- Status Updates
  - Test Specifications (LSE-419)
  - Example notebooks
  - Bootcamps
- Planned Work for Upcoming Year
- Science Validation Surveys

SV Technical Scope and Requirements

1. Determining whether the specifications defined in the ***SRD*** (LPM-17) and ***LSR*** (LSE-29) can be met with the full survey
2. Characterizing ***other system performance metrics*** in the context of the four primary science drivers
3. Studying ***environmental dependencies*** and ***technical optimization*** that inform early operations
4. ***Documenting*** system performance and verifying mechanisms to ***monitor*** system performance during operations
5. Validating ***data delivery***, derived ***data products***, and ***data access*** tools that will be used by the science community

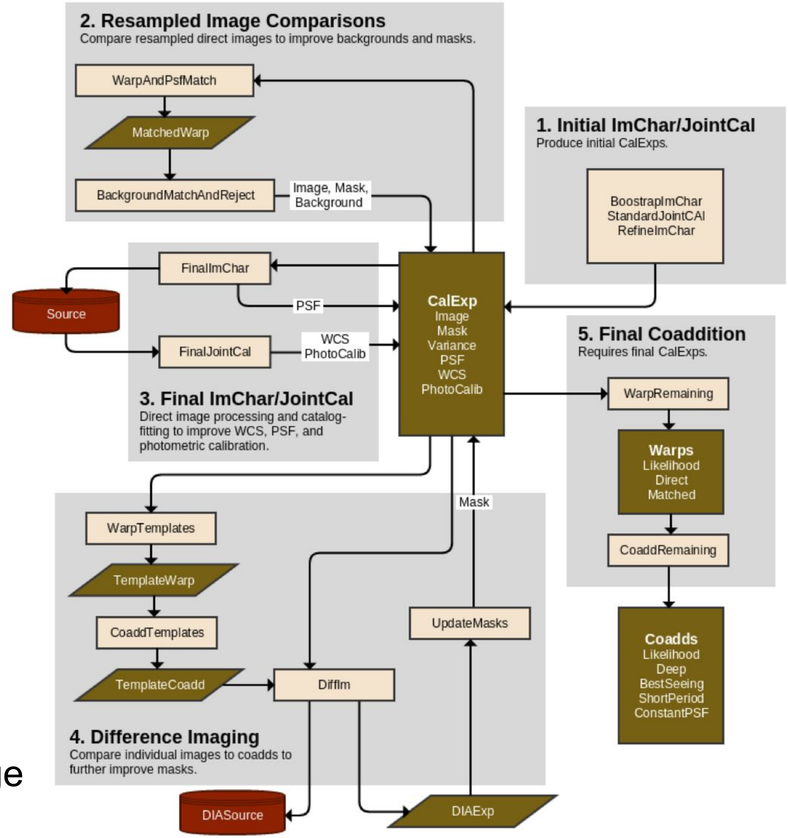
*Aim to quantify the **range of demonstrated performance** by using a combination of on-sky data, informed simulations, and external datasets*

(Re-)Verify Science Pipelines

Science pipelines will have been extensively tested with pre-cursor datasets and LSST simulations as part of DM construction

We will re-verify pipeline components (LDM-151) with data from as-built system:

- 18 calibration products
- 14 APP pipeline components
- 26 DRP pipeline components

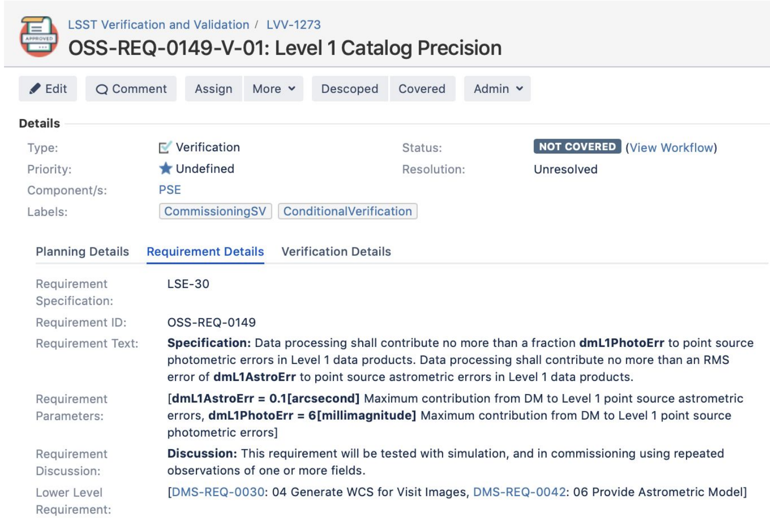


Example:
Data Release Processing image coaddition and differencing

Status: Verification Architecture

Commissioning Science Verification and Validation

- Implemented framework for developing and tracking test cases for OSS and LSR requirements utilizing the JIRA based LSST Verification Architecture
- Automated generation of LSE-419 (Commissioning Science Verification Test Specification Document) from JIRA elements
- Integrated Commissioning SV Test Planning with LDM-639 (DM acceptance test specifications) and LSE-61 (DMSR)


LSST Verification and Validation / LVV-1273

OSS-REQ-0149-V-01: Level 1 Catalog Precision

Edit | Comment | Assign | More | Descoped | Covered | Admin

Details

Type:	Verification	Status:	NOT COVERED (View Workflow)
Priority:	Undefined	Resolution:	Unresolved
Component/s:	PSE		
Labels:	CommissioningSV ConditionalVerification		

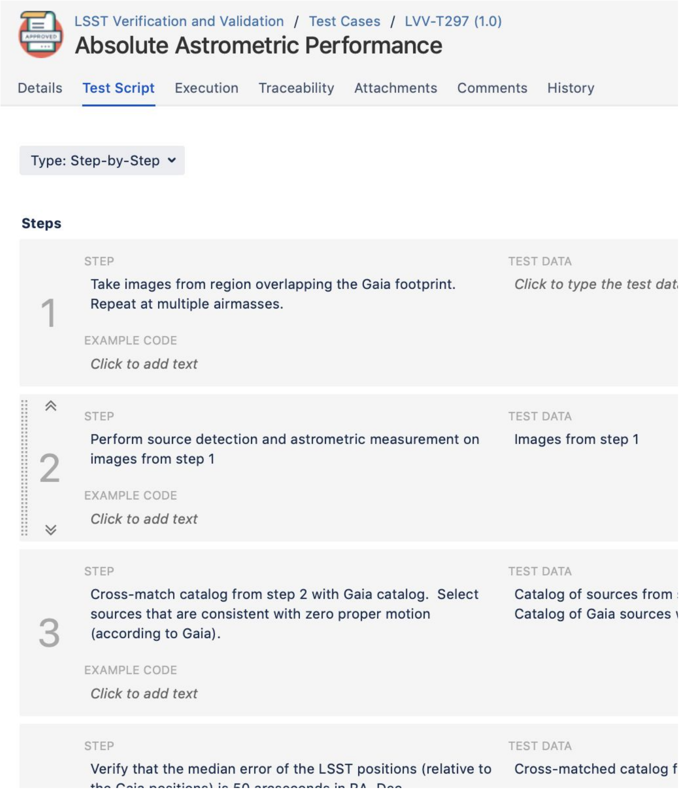
Planning Details | Requirement Details | Verification Details

Requirement Specification:	LSE-30
Requirement ID:	OSS-REQ-0149
Requirement Text:	**Specification:** Data processing shall contribute no more than a fraction **dmL1PhotoErr** to point source photometric errors in Level 1 data products. Data processing shall contribute no more than an RMS error of **dmL1AstroErr** to point source astrometric errors in Level 1 data products.
Requirement Parameters:	**[dmL1AstroErr = 0.1[arcsecond]** Maximum contribution from DM to Level 1 point source astrometric errors, **dmL1PhotoErr = 6[millimagnitude]** Maximum contribution from DM to Level 1 point source photometric errors]
Requirement Discussion:	**Discussion:** This requirement will be tested with simulation, and in commissioning using repeated observations of one or more fields.
Lower Level Requirement:	[DMS-REQ-0030: 04 Generate WCS for Visit Images, DMS-REQ-0042: 06 Provide Astrometric Model]

Status: Test Case Specifications

Test Case Development

- Initial draft test cases for 49/52 OSS and 10/18 LSR requirements related to high-level science performance
- Initial implementation of Jupyter notebooks for developing test cases using precursor data
- Creation of continuous testing environment for Jupyter notebooks using Github



Example Notebook: [OSS-REQ-0388]()

LVV-T297: Absolute Astrometric Performance

Written By: Bryce Kalmbach

Last updated: 07-10-2019

Tested on Stack Version: w_2019_27

Requirements:

OSS-REQ-0388

Median error in absolute position for each axis, RA and DEC, shall be less than 50 milliarcseconds.

Proposed Test Case:

1. Take images from region overlapping the Gaia footprint. Repeat at multiple airmasses.
2. Perform source detection and astrometric measurement on images from step 1
3. Cross-match catalog from step 2 with Gaia catalog. Select sources that are consistent with zero proper motion (according to Gaia).
4. Verify that the median error of the LSST positions (relative to the Gaia positions) is **50 milliarcseconds in RA, Dec independently**

Import necessary tools

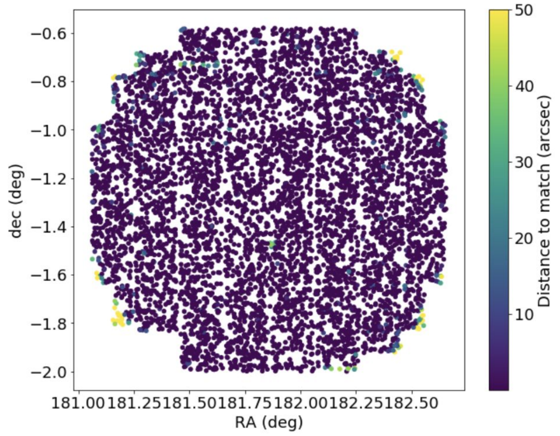
```
In [ ]: import os
        import numpy as np
        import matplotlib as mpl
        import matplotlib.pyplot as plt
        import pandas as pd
```

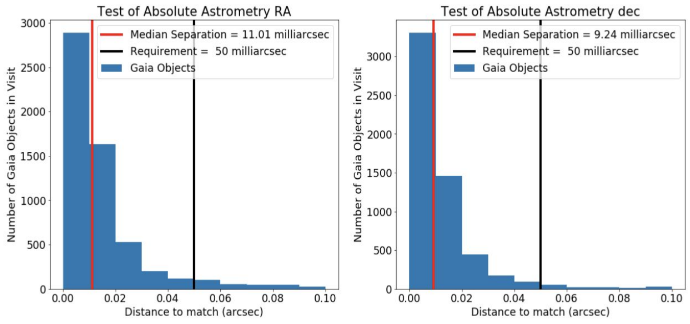
```
In [ ]: from lsst.daf.persistence import Butler
        import lsst.daf.persistence as daf_persistence

        from astropy.coordinates import SkyCoord
        from astropy import units as u
```

```
In [ ]: # Make our plots nice and readable
        plt.rcParams.update({'font.size': 18})
```

Identify HSC Data to use

We want to get data from a single visit for this requirement so we choose a visit from the HSC Wide dataset. https://hsc-release.mtk.nao.ac.jp/doc/index.php/database/ has info on which tracts are included in the Wide data. We randomly choose





The requirements are satisfied if both RA and dec median values are less than 50 milliarcseconds.

Status: Training and Development

Data Management, Camera, Commissioning bootcamp 13-16 Nov 2018 (SLAC)

- 35 attendees with representation from Commissioning, Camera, and Data Management subsystems
- Focused on instrument signature removal

Science Verification Test Specifications bootcamp 10-12 June 2019 (Tucson)

- 20 attendees with representation from Commissioning, Camera, T&S, and Data Management subsystems
- Focused on training in the verification architecture, review of draft test cases for OSS and LSR requirements, and the implementation of these test plans using precursor and simulated data

Personnel

Data Release Processing:

Keith Bechtol - University of Wisconsin

Chris Walter - Duke University

Tony Tyson - UC Davis

Sam Schmidt UC Davis

Andrew Bradshaw - UC Davis

Imran Hassan - UC Davis

Jim Bosch - Princeton University

Yusra AlSayyad - Princeton University

Sophie Reed - Princeton University

Nate Lust - Princeton University

Dan Taranu - Princeton University

C. Waters - Princeton University

Alert Production Processing:

Andrew Connolly - University of Washington

Bryce Kalmbach - University of Washington

Scott Daniel - University of Washington

Meredith Rawls - University of Washington

Eric Bellm - University of Washington

Mario Juric - University of Washington

Eve Kovacs - Argonne National Lab

Ian Sullivan - University of Washington

Italics = Support assigned from Data Management Team

Management and Organization

Calibration Products Processing:

Merlin Fischer-Levine - Princeton University

Christopher Stubbs - Harvard University

Patrick Ingraham - AURA - LSST

Robert Lupton - Princeton University

August Guyonnet - Harvard University

Italics = Support assigned from Data Management Team

Analysis of commissioning data products is intrinsically a test of both the hardware performance as well as the science pipelines and data access tools

Single Commissioning Science Validation effort coordinated with Data Management construction effort

Management and Organization

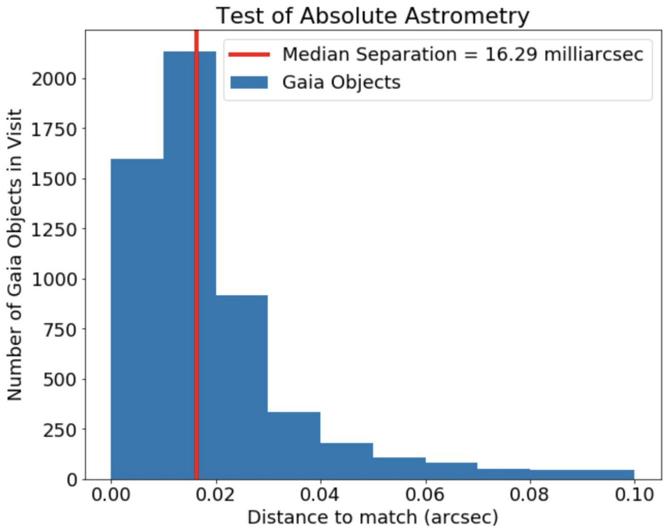
Additional Support:

- Following core AI&T activities, some members of System Integration Team are planned to transition to science validation activities (e.g., Brian Stalder, Sandrine Thomas)
- Commissioning budget includes resources to enlist topical experts from the broader science community for specific analysis tasks (sabbatical support)
- 20% of DM construction effort during commissioning is set aside for responding to algorithmic or data discoveries (part of DM construction budget)

Upcoming Year: Getting Ready for Data

Implementation of Test Cases using Precursor Data and Simulations

- Development and documentation of OSS and LSR test cases using Jupyter notebooks and existing data sets (e.g. HSC and DECam)
- Implementation of test cases within the DM (SQuaSH) framework (automated evaluation of performance metrics) to track metrics against data set and code revision



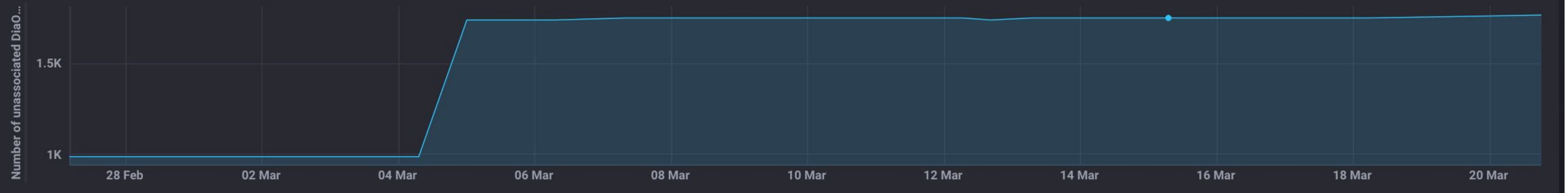
Analysis of Site-specific Data

- For example, analysis of DIMM data from the site using DM tools to evaluate image quality as a function of time

Definition of Requirements for Commissioning Verification Surveys

SQuaSH Metrics Dashboard

Total Unassociated DiaObjects



Total Unassociated DiaObjects

Time (UTC)	RUN	totalUnassocObjs	Code Changes
03/20/2019 18:28:49	175	1772	afw, meas_modelfit, meas_extensions_astrometryNet, obs_lsst, jointcal, lsst_dm_stack_demo, pex_config, verify, ap_verify, ip_
03/18/2019 06:20:24	172	1756	
03/17/2019 06:24:43	171	1756	
03/16/2019 06:19:22	170	1756	
03/15/2019 07:04:26	169	1756	jointcal, obs_base
03/14/2019 07:09:08	168	1756	daf_butler

Planning Tests of Increasing Sophistication

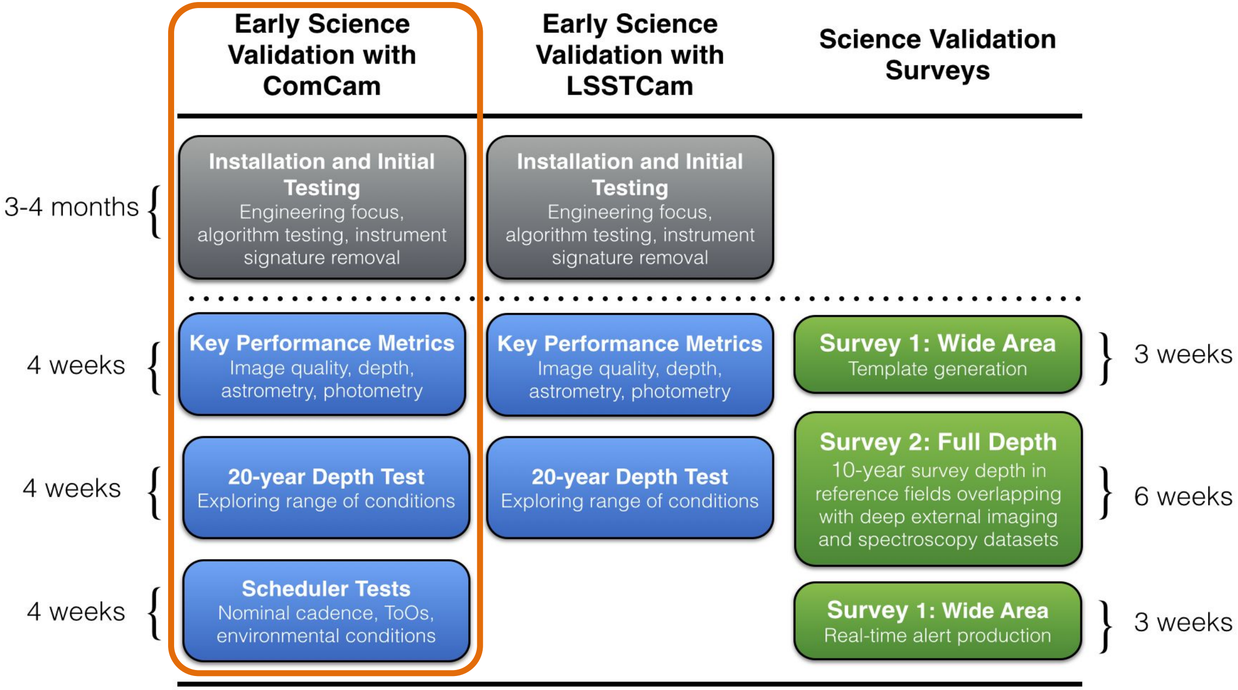
- Verify with on-sky data as early as possible
- Gradual transition from engineering activities to sustained operations
  - Engineering focus during AI&T with ComCam and LSSTCam
  - Allocate ~25% of total time for engineering activities during early Science with ComCam and LSSTCam
  - Approach early operations level during Science Validation Surveys
- Tests of increasing sophistication: calibration products → single-visit performance → image stack performance → other metrics
- Direct test if possible; validate with simulations otherwise
  - Simulations used to assess expected 10-yr proper motion precision, 10-year survey coverage, detection completeness

Planned On-sky Observing Campaigns

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

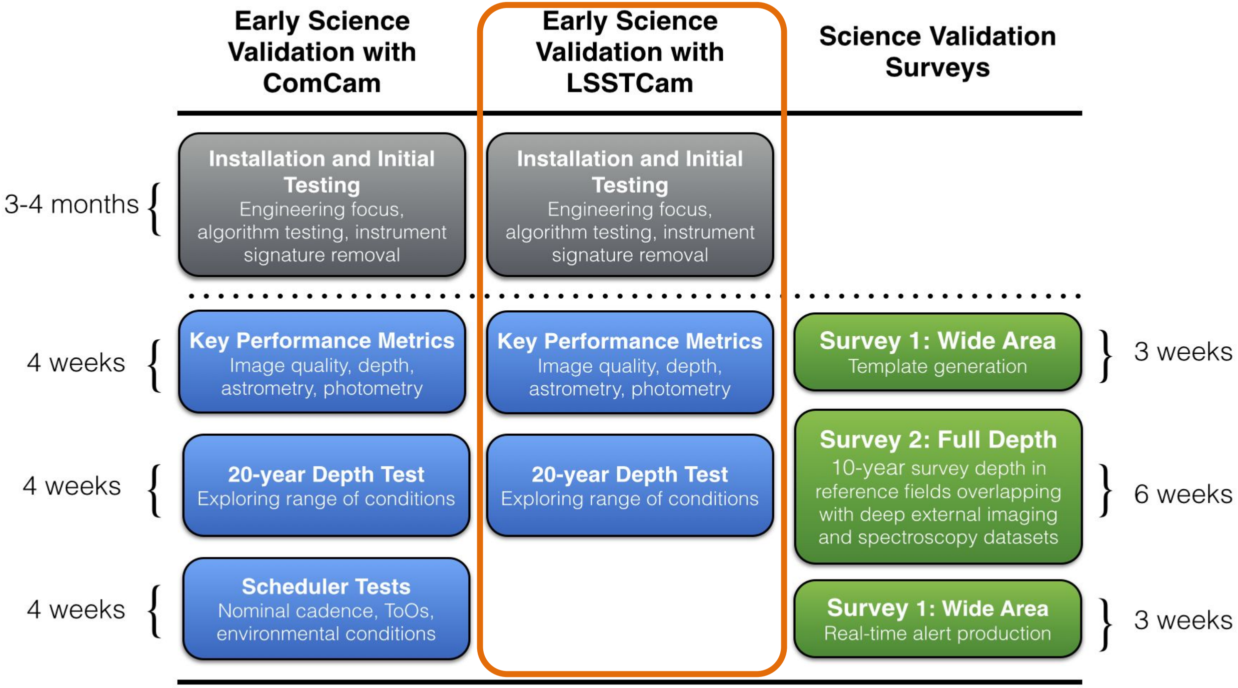
Early Science Validation with ComCam

- Science images with ComCam provide a first opportunity for many tests
- Repeated imaging of several fields in multiple bands at different airmasses, source densities, etc.
- Exploring range of environmental conditions
- Scheduler testing in variety of observation modes with actual telemetry



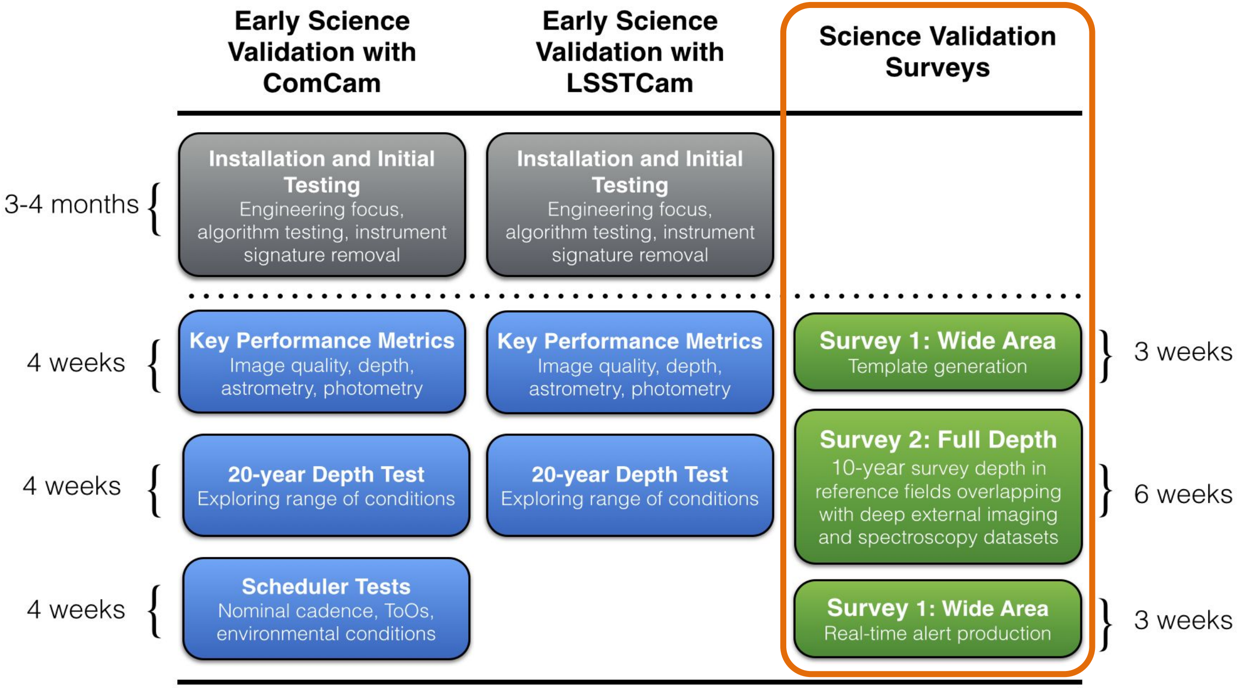
Early Science Validation with LSSTCam

- Repeat sequence of early science verification observations and analysis from ComCam with LSSTCam, making use of experience and analysis tools gained with ComCam
- Focus on range of delivered performance over larger FOV



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



Additional Slides

JIRA Test Specification Planning

OSS-REQ-0149: Level 1 Catalog Precision

LSST Verification and Validation / LVV-1273

1 of 2
Return to search

OSS-REQ-0149-V-01: Level 1 Catalog Precision

Edit | Comment | Assign | More | Descoped | Covered | Admin | Email | Pivot Report | Export

Details

Type:	Verification	Status:	NOT COVERED (View Workflow)
Priority:	Undefined	Resolution:	Unresolved
Component/s:	PSE		
Labels:	CommissioningSV ConditionalVerification		

Planning Details | Requirement Details | Verification Details

Requirement Specification:	LSE-30
Requirement ID:	OSS-REQ-0149
Requirement Text:	**Specification:** Data processing shall contribute no more than a fraction **dmL1PhotoErr** to point source photometric errors in Level 1 data products. Data processing shall contribute no more than an RMS error of **dmL1AstroErr** to point source astrometric errors in Level 1 data products.
Requirement Parameters:	[**dmL1AstroErr = 0.1[arcsecond]** Maximum contribution from DM to Level 1 point source astrometric errors, **dmL1PhotoErr = 6[millimagnitude]** Maximum contribution from DM to Level 1 point source photometric errors]
Requirement Discussion:	**Discussion:** This requirement will be tested with simulation, and in commissioning using repeated observations of one or more fields.
Lower Level Requirement:	[DMS-REQ-0030: 04 Generate WCS for Visit Images, DMS-REQ-0042: 06 Provide Astrometric Model]

People

Assignee:	Scott Daniel Assign to me
Reporter:	Syndeia PSE User
Votes:	0 Vote for this issue
Watchers:	4 Stop watching this issue

Dates

Created:	19/Aug/18 2:49 AM
Updated:	1 minute ago
Resolved:	19/Aug/18 2:49 AM

CI Builds

No builds found.

Agile

View on Board

Test Case Flowdown for Requirements

Description

This sub-requirement will focus on the astrometric precision part of OSS-REQ-0149.

Note: given that it is very hard to separate contributions due to software and hardware and because the astrometric requirements in OSS-REQ-0388 (LVV-1363) are more stringent, we are marking this requirement as "ConditionalVerification" meaning, as long as OSS-REQ-0388 passes, this requirement will have been met. If OSS-REQ-0338 fails, it will be necessary to try to determine why (i.e. was it due to observing conditions or the DM software).

Actual test procedure:

Insomuch as this requirement is going to be tested with the repeatability of pairwise astrometric measurements, it will be complementary with LSR-REQ-0094 (LVV-238).

Summary of discussion on May 13, 2019:

– Verifying the repeatability of pairwise astrometric measurements will test this requirement, since the the only variation in those measurements should be that which is introduced by the Data Management processing pipeline.

– Measuring the RMS of the astrometric solution with respect to Gaia is probably also good enough to test this requirement.

– While running simulations with the atmosphere "turned off" would allow us to isolate only those contributions to astrometric uncertainty that originate in the DM pipeline, it is probably more effort than it is worth to generate and process those simulations.

Traceability

Test Cases

Coverage

› LVV-T297 (1.0) Absolute Astrometric Performance	DEFINED
› LVV-T545 (1.0) Astrometric error -- level 1 processing -- reference catalog	DEFINED
› LVV-T959 (1.0) Inter-band astrometric consistency	DEFINED
› LVV-T960 (1.0) Relative astrometric performance	DEFINED

Test Case Specification and Tracking

LSST Verification and Validation / Test Cases / LVV-T297 (1.0)

Absolute Astrometric Performance

Back | **Save** | New Version | 1.0 ···

Details | **Test Script** | Execution | Traceability | Attachments | Comments | History

Type: Step-by-Step

Steps

#	STEP	TEST DATA	EXPECTED RESULT
1	Take images from region overlapping the Gaia footprint. Repeat at multiple airmasses. EXAMPLE CODE: *Click to add text*	*Click to type the test data*	Set of images
2	Perform source detection and astrometric measurement on images from step 1 EXAMPLE CODE: *Click to add text*	Images from step 1	Catalog of sources
3	Cross-match catalog from step 2 with Gaia catalog. Select sources that are consistent with zero proper motion (according to Gaia). EXAMPLE CODE: *Click to add text*	Catalog of sources from step 2 Catalog of Gaia sources with measured positions	Cross-matched catalog of sources seen by both LSST and Gaia
	Verify that the median error of the LSST positions (relative to	Cross-matched catalog from step 3	*Click to type the expected result*

Add step · Call to test · Clone · Attach files · Delete

Test Case Specification Dashboard

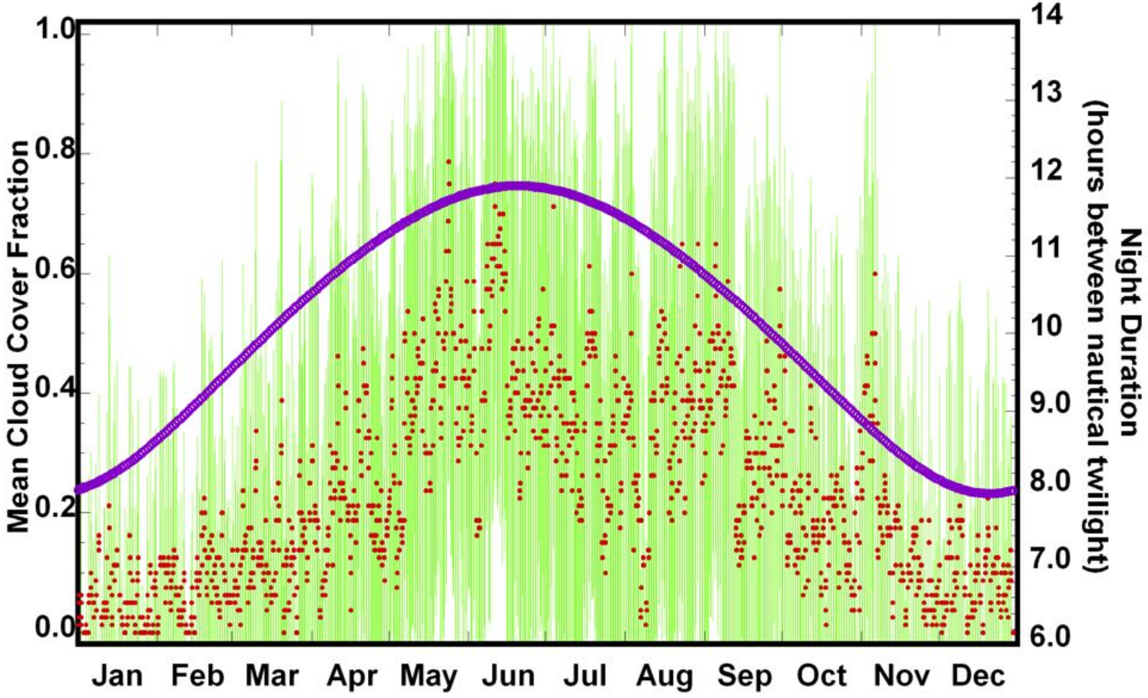
2. Status Summary

	OSS	LSR
Summary Coverage Report	link	link
Total Number of Verification Elements in Commissioning SV	134 issues	84 issues
Verification Elements that have test cases defined and ready to be checked	73 issues	16 issues
Verification Elements with "Delete" label	0 issues	0 issues
Verification Elements that have no associated test case	53 issues	66 issues
Requirements with no associated test case	3 issues	18 issues
Verification Elements that have test cases not in the /Project Systems Engineering/Commissioning Science Verification|LSE-419 folder	4 issues	2 issues
Verification Elements that require more thought on the specification	17 issues	27 issues

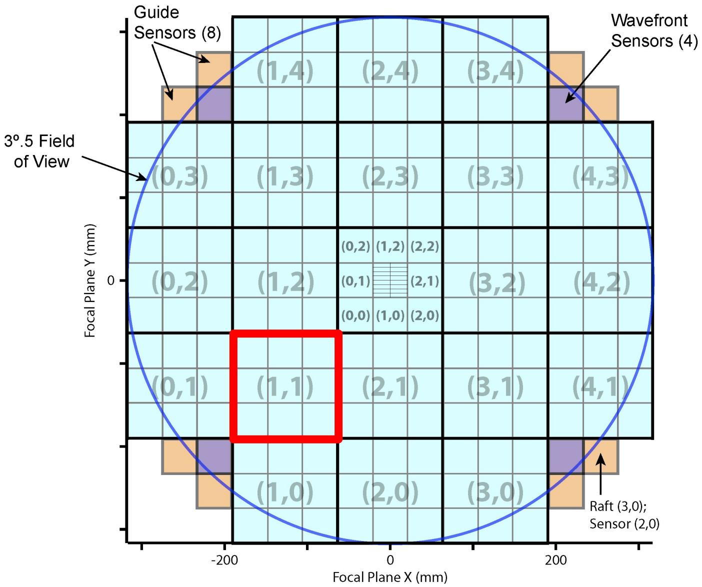
On-Sky Observing Campaigns

Taking Weather Into Account

When planning the time needed for on-sky observations, we have assumed that (on-average) 85% of time is usable and 53% of time is photometric. Historical weather patterns at CTIO suggest that the number of hours of dark clear skies per night (~8) is approximately constant over the annual cycle.



Focal Plane Size, Expected Source Counts



Raft area (ComCam) ~ 1600 $arcmin^2$
~ 0.45 deg^2

Full LSST camera area ~ 9.6 deg^2

Sample (typical high Galactic latitude field)	Density ($arcmin^{-2}$)	# Per ComCam FOV	# Per LSSTCam FOV
High SNR stars useful for PSF determination	~3	~5K	~100K
“Gold” sample of galaxies	~55	~90K	~2M
Galaxies useful for weak lensing	~40	~60K	~1.4M

Planned On-sky Observing Campaigns

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

ComCam AI&T On-Sky Observations

Objectives

- Focus on electro-optical tests, engineering, instrument signature removal
- First on-sky data

Example observations

- Build and test pointing model
- Build and test active optics system look-up table, wave front sensors
- Raster single field across each detector to determine illumination corrections, initial color-term, and verify astrometric solutions (star flats)
- Repeated observations to test stability of photometric and astrometric solutions and statistical precision
- Repeated observations of celestial pole at different rotations (fixed airmass effects)
- Observations of celestial pole through different amounts and kinds of clouds

ComCam KPM Testing

Objectives

- Evaluate Key Performance Metrics (KPMs) for single-visit performance (e.g., relative + absolute photometry and astrometry, image quality, throughput)
- Measure residual PSF ellipticity distribution; test transient and moving object detection + linkage

Observations

- 20 fields x 5 epochs x 5 visits x 6 filters = 3K visits (~4 nights)
- Several fields contain absolute photometric calibration standards
- Range of airmass, source densities
- 3 fields x 3 (dither allowance) x 200 visits x 2 filters (r, i) = 3.6K visits (~5 nights)
- Sample range of source densities, at least one along ecliptic

ComCam 20-year Depth Testing

Objectives

- Focus on image stack performance, sampling range of conditions
- Identify subsets of the data for Data Release Processing (e.g., best/worst seeing, lowest/highest airmass)
- Repeated observations of the same fields are useful for testing template generation algorithms and Alert Processing pipelines (can be offline)

Observations

- Observe 10 fields to depth equivalent to 20 years of Wide-Fast-Deep survey in 6 filters (~1700 visits per field, ~20 nights)
- Where possible, fields should overlap external reference datasets
- Explore a range of environmental conditions to examine various potential systematics — observations driven by needs to test pipeline algorithms
- Dither pointings in each field to approximate Wide-Fast-Deep pattern

ComCam Scheduler Testing

Objectives

- Validate predictions of operations simulator
- Test scheduler feedback with real telemetry (including auxiliary instruments)
- Exercise interfaces and procedures used by human operators during normal operations
- Measurements of slew and settle times with realistic observing patterns

Observations

- Run automated scheduler with normal cadence under range of environmental conditions
- Testing special observation modes, e.g., Target-of-Opportunity interrupts, survey over constrained area, modified tactician
- Observations may be interspersed with 20-year depth test

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

- ~1600 deg2 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

- ~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