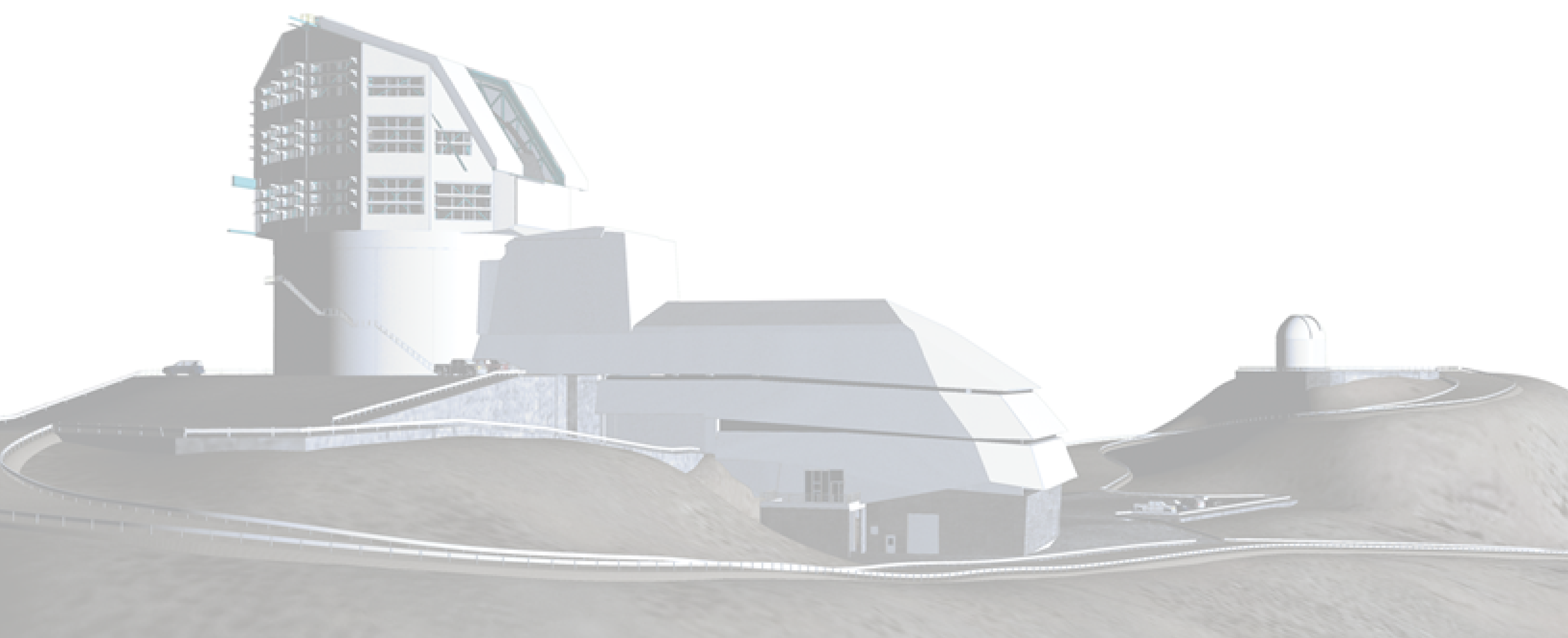
Vera C. Rubin Observatory
Data Management

Data Management for LSST Special Programs

M. L. Graham, Y. AlSayyad, K. Bechtol, E. C. Bellm, J. F. Bosch, J. L. Carlin, G. P. Dubois-Felsmann, L. P. Guy, M. Jurić, K.-T. Lim, E. S. Rykoff, C. T. Slater, and the Data Management System Science Team

DMTN-065

Latest Revision: 2024-03-01



Abstract

Special Programs are additional survey areas and/or observing strategies that are driven by specific science goals which build on, or are beyond, the core science pillars of the Wide Fast Deep Main Survey. In order to meet the requirements and enable science related to Special Programs, this document provides recommendations for Rubin Data Management regarding processing and serving data products for Special Programs.

Hardware and processing boundaries on the potential diversity of data from Special Programs are discussed along with scenarios in which user-generated processing and data products might be needed to meet Special Programs' science goals.

Change Record

Version	Date	Description	Owner name
0	2017-11-14	Status: internal working document.	Melissa Graham
1	2018-06-17	Updated to finalize and issue.	Melissa Graham
2	2021-12-01	Updates per DM-20375.	Melissa Graham
3	2023-10-19	Updates per DM-32723.	Melissa Graham

Document source location: `https://github.com/lsst-dm/dmtn-065`

Contents

1 Memoranda **1**

2 Terms and Definitions **1**

2.1 Visit Types 1

2.2 Region Label 2

2.3 Observing Mode Label 2

2.4 Main Survey including Wide-Fast-Deep (WFD) 2

2.5 Special Programs 3

2.6 Special Processing 3

2.7 Standard (non-special) processing 4

2.8 User-Generated Processing 4

2.9 Deep Drilling Field (DDF) 4

2.10 Mini-, Micro- and Nano-Surveys 5

3 Recommendations to Enable Science from Special Programs **5**

3.1 Region and Observing Mode Labels 5

3.2 Special Processing 7

3.3 Prompt Processing 8

3.4 Computational Resources for User-Generated Processing 8

3.5 Rubin Science Platform Capabilities 9

4 Details and Challenges of Processing Data from Special Programs **9**

4.1 Boundaries on Non-Standard Visits 9

4.1.1 Hardware Boundaries 9

4.1.2 Processing Boundaries 10

4.2 Standard (non-special) processing 11

4.2.1 Prompt Processing and Alert Production 11

4.2.2 Solar System Processing 13

4.2.3 Data Release Processing 13

4.3 Special Processing . . . 14
4.3.1 Scenarios to illustrate “possible and necessary” . . . 14
4.3.2 Deep Drilling Fields (DDFs) . . . 15
4.3.3 Short-Exposure Mini-Surveys . . . 16
4.3.4 Standard Visit Mini-Surveys . . . 16
4.3.5 Target-of-Opportunity (TOO) Observations . . . 17
4.4 User-Generated Processing . . . 17
4.4.1 Deep Drilling Fields . . . 18
4.4.2 Short-Exposure Mini-Surveys . . . 18
4.4.3 Mini-Surveys . . . 18

5 Requirements Related to Special Programs 18
5.1 Metadata . . . 18
5.2 Data Products . . . 19
5.3 Processing . . . 20

A Hypothetical Scenarios 22
A.1 Outer solar system mini-survey . . . 22
A.2 Deep Drilling Field . . . 24
A.3 Short-exposure twilight survey . . . 25

B Potential Hardware Boundaries 25
B.1 Filter Changes . . . 26
B.2 Filter Carousel Loads . . . 26
B.3 Exposure Times . . . 26
B.4 Readout Time . . . 26
B.5 Inter-Image Time . . . 27
B.6 Telescope Slew . . . 27
B.7 Pointing and Sidereal Tracking . . . 27
B.8 Non-Sidereal Tracking . . . 28
B.9 Camera Rotation . . . 28

C Potential Processing Boundaries **28**
C.1 Exposure Times 29
C.1.1 Short Exposures (Non-Standard Visits of <30 sec) 29
C.1.2 Long Exposures (Non-Standard Visits of >30 sec) 31
C.2 Twilight Images with a Bright Background 31
C.3 Images Obtained with Non-Sidereal Tracking 31

D Documentation Review for Requirements Related to Special Programs **32**
D.1 Science Requirements Document (SRD) 32
D.2 LSST System Requirements (LSR) 32
D.3 Observatory System Specifications (OSS) 34
D.4 Data Management Subsystems Requirements (DMSR) 36
D.5 Data Management Applications Design (DMAD) 38
D.6 Data Products Definitions Document (DPDD) 39

Data Management for LSST Special Programs

1 Memoranda

The three memoranda for Rubin Data Management regarding Special Programs are:

1. All visits should have a region label *and* an observing mode label, and these labels should be propagated to processed images and catalog data. (Section 3.1)
2. Special Processing should be done by Rubin Data Management to produce unique and separate data products for Special Programs when it is both possible and necessary. (Section 3.2).
3. All visits that *can* be processed by the Prompt pipelines and generate alerts *should* be, in support of time domain and Solar System science goals. (Section 3.3)

2 Terms and Definitions

The following is an introduction to the key terms related to Special Programs.

2.1 Visit Types

A visit is an observation of a single pointing at a given time, of which there are three types:

- Standard Visit – Composed of $2{\times}15$ second exposures (commonly referred to as "snaps").
- Alternative Standard Visit – Composed of a single 30 second exposure.
- Non-Standard Visit – Any other exposure time(s) or number of snaps.

Non-standard visits with shorter or longer exposure times are being considered for some Special Programs.

2.2 Region Label

Based on Right Acension and Declination.

To be applied to data by Rubin Data Management, e.g., using a look-up table.

Region labels would include, e.g., WFD low-dust, Galactic Plane, Galactic Bulge, SCP, NES, or Virgo; DDF fieldname; mini-survey region name.

2.3 Observing Mode Label

Based on the scheduler mode at the time of the observation (i.e., survey cadence or strategy).

To be applied by the Survey Scheduler team via the observations scheduler.

Observing mode labels might include, e.g., WFD, DDF, TOO, mini-survey name, or other options like engineering, commissioning, and director's discretionary.

2.4 Main Survey including Wide-Fast-Deep (WFD)

The LSST Main Survey includes the low-dust Wide-Fast-Deep (WFD) area (which includes low-dust extragalactic and Galactic regions) and several special regions: the dusty areas of the Galactic bulge and plane; the South Celestial Pole (SCP); the North Ecliptic Spur (NES); and the Virgo Cluster [PSTN-055]. The WFD component is the core component of the LSST Main Survey, designed to achieve the science goals defined by the Science Requirements Document (SRD; LPM-17).

The Main Survey will be executed with a set of observing modes, and is expected to use alternative standard visits in the u-band and standard visits in all other bands [PSTN-055]. Different cadences, filter balances, rotational dithers, etc., might be applied in the different regions composing Main Survey. For example, rolling cadence is expected to be implemented in the low-dust WFD region only.

Data from the Main Survey will be processed with standard (non-special) processing (Section 2.7). This processing will produce, among other data products (LSE-163), a contiguous sky footprint that covers at least $\sim$18000 deg^2 (with $\gtrsim$800 visits per field in WFD regions), and is expected to be accomplished with 85–90% of the observing time. This contiguous all-sky ”deep coadd” may have region-dependent processing parameters and inputs (e.g., different calibration parameters or deblending algorithms for crowded fields) and will be of variable depth, as most special regions recieve fewer visits.

2.5 Special Programs

This is a Rubin Data Management term used to refer to sky regions within or beyond the Main Survey footprint that have a distinct observing strategy from the Main Survey's plans for that region (e.g., different visit type, cadence, filter distribution).

Special Programs are typically driven by specific science goals that build on or add to the core science pillars of the LSST. They include LSST components such as the Deep Drilling Fields (DDFs) and the mini-, micro-, and nano-surveys [PSTN-055]. About 10–15% of the total 10-year LSST will be spent obtaining observations associated with Special Programs.

Data products from Special Programs is subject to the Rubin Data Policy [RDO-013] in the same way as data products from the Main Survey.

Some science goals for Special Programs can be met with standard processing, but some will require Special Processing by Rubin Data Management or user-generated processing.

2.6 Special Processing

This is a Rubin Data Management term to describe processing that uses components of the LSST Science Pipelines and is applied by Rubin Data Management to images from Special Programs. Special Processing creates data products that are unique and separate from those produced by standard processing for the Main Survey.

Special Processing is likely to use different inputs or configurations for the LSST Science Pipelines, or to run on different timescales, than standard processing - as appropriate for the Special Programs' data and science goals. However, the development and application of specialized *algorithms* or new software is beyond the scope of Special Processing.

Examples of Special Processing include the processing of images with non-standard exposure times, images obtained during twilight, or nightly stacking and differencing of images in deep drilling fields. See Section 4.3 for more detailed examples of Special Processing.

2.7 Standard (non-special) processing

This is a term used only in this document to refer to the image processing described in the Data Management Science Pipelines Design LDM-151 that produces the data products described in the Data Products Definitions Document (LSE-163) that are designed for, and will be applied to, the Main Survey's observations.

In some cases, standard processing is also appropriate for Special Programs.

2.8 User-Generated Processing

Any processing of Rubin data done by users in order to reach specific science goals, including processing for Special Programs data, is referred to as User-Generated Processing.

User-Generated Processing for Special Programs data would be necessary in cases where the science goals require custom algorithms, software, or very large computational capacities which are beyond the scope of Special Processing or the Rubin-provided computational resources (Section 3.4).

Guidelines for User-Generated Processing, and for user-generated data products that can be federated with the Rubin-product data products (i.e., joinable tables), is forthcoming.

2.9 Deep Drilling Field (DDF)

A single pointing for which many (e.g., a hundred) visits are obtained (usually sequentially) during a single night, and repeated every few nights.

As of the Phase 2 SCOC recommendations in PSTN-055, the five confirmed DDFs were:

- Elias S1 (00:37:48, -44:00:00)

- XMM-LSS (02:22:50, -04:45:00)
- Extended Chandra Deep Field-South (03:32:30, -28:06:00)
- COSMOS (10:00:24, +02:10:55)
- Euclid Deep Field South (04:04:58, -48:25:23)[1]

Processing for the DDF images is likely to be a combination of standard (Prompt) processing, Special Processing, and User-Generated Processing.

2.10 Mini-, Micro- and Nano-Surveys

Specific sky areas covered by a few hundred, a hundred, or tens of visits (respectively). This document will refer to them collectively as mini-surveys.

The sky areas of mini-surveys can be within, adjacent to, or detached from the Main Survey footprint. Mini-surveys can have non-standard visits. Target-of-opportunity (TOO) observations for, e.g., the discovery of optical counterparts to multi-messenger astrophysical phenomena, are considered a type of mini-survey in this document. For a list of the mini-surveys under consideration, see PSTN-055.

Processing for the mini-surveys is likely to be a combination of standard (Prompt) processing, Special Processing, and User-Generated Processing.

3 Recommendations to Enable Science from Special Programs

Recommended implementations for Rubin Data Management to meet the requirements related to Special Programs (Section 5) and enable science with data from Special Programs.

3.1 Region and Observing Mode Labels

All visits should have at least one region label *and* one observing mode label, and these labels should be propagated to processed images and catalog data.

[1] https://www.cosmos.esa.int/web/euclid/euclid-survey

Region Labels – Based on Right Acension and Declination. To be applied to data by Rubin Data Management, e.g., using a look-up table. Region labels would include, e.g., WFD low-dust, Galactic Plane, Galactic Bulge, SCP, NES, or Virgo; DDF fieldname; mini-survey region name. Region labels should be propagated to processed images (visit, difference, and deep coadds), catalog sources, and catalog static-sky objects. A given image might have more than one region label due to overlaping regions, but they could be, e.g., held in a single string as a CSV.

Observing Mode Labels – Based on the scheduler mode at the time of the observation (i.e., survey cadence or strategy). To be applied by the Survey Scheduler team via the observations scheduler. Observing mode labels might include, e.g., WFD, DDF, TOO, mini-survey name, or other options like engineering, commissioning, and director's discretionary. Observing mode labels should be propagated to single-visit processed and difference images and catalog sources. It would not make sense to propagate observing mode labels to deep coadds or object tables, as these data products could be a mix of modes. It is left to the discretion of the Survey Scheduler team whether more than one observing mode labels might be applied at a time and, e.g., held in a single string as a CSV.

A given image would probably only have one observing mode label

There are three main motivations for this recommendation.

1. To meet the requirement that metadata for observations associated with Special Programs is stored, and is sufficient for triggering real-time data processing recipes (Section 5.1).
2. To enable users to query and retrieve processed image and catalog data associated with a specific Special Program, and meet the science goals of that Special Program, when standard processing as been applied (e.g., Prompt Processing, Section 3.3).
3. To enable provenance when Special Programs data in included in standard processing and the Main Survey's data products (e.g., if used to improve the all-sky coadd).

3.2 Special Processing

Special Processing should be done by Rubin Data Management to produce unique and separate (and joinable) data products for Special Programs when it is both possible and necessary.

Possible – When original or reconfigured versions of the LSST Science Pipelines can be used, and no new algorithmic or software development, or significant additional computational resources, are needed.

Necessary – When the primary science goal for a Special Program cannot be met by including the data in standard processing (e.g., Prompt processing), or where doing so would compromise the Main Survey data products (e.g., introduce additional non-uniformity beyond what would be expected based on the survey strategy).

There are two main motivations for this recommendation.

1. To meet the requirement that Rubin Observatory produce unique, separate, and joinable data products whenever this is possible with the original or reconfigured versions of the LSST Science Pipelines (Section 5.2).
2. To enable science with Special Programs by all users, not just those with the time and effort to process the data, and to reduce computational load (and potential redundancy) in User-Generated Processing.

The definitions of possible and necessary are further illustrated with examples in Section 4.3. Note that secondary science goals may be considered as "not necessary".

The general scope of Special Processing, including situations in which cross-match and table joins would be possible and scientifically relevant, are ultimately left to the discretion of Rubin Data Management in Operations.

Further discussion on Special Processing, with examples, is provided in Section 4.3.

3.3 Prompt Processing

All visits that *can* be processed by the Prompt pipelines and generate alerts *should* be, in support of time domain and Solar System science goals.

The condition "can be processed" is ultimately left to the discretion of Rubin Data Management in Rubin Operations, but it is expected to include all standard and alternative visits in sky regions for which a template image exists.

The main motivation for this recommendation is that all time-domain and moving-object science goals (two of the four science pillars for the LSST) are enhanced by any and all additional observations, even if they are not optimized for the science (as the WFD area's cadence is being optimized).

Further discussion of standard Prompt processing of Special Programs data, with examples and a discussion of a few challenges, is provided in Section 4.2.1.

3.4 Computational Resources for User-Generated Processing

Rubin Observatory will reserve 10% of its total data processing capacity for users. This component would include *all* user processing and re-processing of any and all LSST data, including Special Programs. This is already a requirement, as mentioned in Section 5.3.

As described in Section 4.4 this processing capacity will be accessed via Rubin Science Platform, with a supported software environment and infrastructure for batch processing [DMTN-202].

Very computationally intense processing (e.g., shift-and-stack for faint moving objects) will likely require external resources[2].

[2] For more details about the boundary between what Rubin Observatory will provide (in terms of data products and processing resources) and what will be left to the expertise of the science community, see `https://www.lsst.org/about/dm/data-products`

3.5 Rubin Science Platform Capabilities

Users will need to be able to query for data that restricts by sky region and observing mode label. This can be accomplished by including those labels in all image and catalog metadata as described in Section 3.1, as the TAP service and butler already provide the mechanism for user-specified queries.

Users will need to be able to discover Special Programs data products when browsing data. This could be accomplished with, e.g., a toggle to overlay region boundaries for Special Programs when browsing all-sky coadds built from Main Survey data in the Portal Aspect.

4 Details and Challenges of Processing Data from Special Programs

A discussion of the anticipated details and challenges related to obtaining and processing data from Special Programs.

Appendix A provides detailed examples for the processing of data from Special Programs, including scenarios in which standard, Special, and User-Generated Processing are all involved.

4.1 Boundaries on Non-Standard Visits

Special Programs that do not use standard or alternative standard visits might be affected by hardware or processing boundaries.

4.1.1 Hardware Boundaries

Appendix B lists all of the hardware boundaries that might constrain the potential diversity of Special Programs data.

In general, the currently-proposed Special Programs in PSTN-055 are not anticipated to be limited by hardware boundaries.

A few potential challenges posed by hardware boundaries are summarized below.

- **Short exposures –** Special Programs that use short exposures would be limited to the minimum exposure time of 1 second (stretch goal: 0.1 seconds[3]). There is a potential hardware boundary that limits the readout rate to 1 every 15 seconds, which would affect the image acquisition rate and increase the overheads on short exposures.
- **Repeated pointing –** Special Programs that require the *exact same* field pointing and rotation for *every exposure* (to sub-arcsecond levels) might run into hardware boundaries on pointing and tracking.
- **Twilight images –** Special Programs that obtain twilight images will be subject to safe limits on sky background flux, as with any astronomical camera.

Finally, as a side note, Special Programs that request a high number of filter changes and/or long slews could be inefficient due to large overheads, but would not be limited by hardware boundaries.

4.1.2 Processing Boundaries

Appendix C describes the boundaries on what types of visits can be processed and calibrated by the Data Management System and the LSST Science Pipelines.

Most of the currently-proposed Special Programs in PSTN-055 are not anticipated to be limited by hardware boundaries. However, those which use non-standard visits, especially those with short exposure time or those obtained during twilight, might be affected by processing boundaries.

The most likely challenges posed by processing boundaries are summarized below.

- **Very short exposures –** Special Programs that use very short (<2 sec) exposures could be difficult to reduce due to an incompletely-formed PSF (Section C.1). The Data Management System is required to be able to process exposure times as low as 1 second (Section 5.3), but it is known that such short exposures might have degraded image quality.

[3] A short-exposure survey of the bright stars of M67, described in Chapter 10.4 of the Observing Strategy White Paper [12], suggests using the stretch goal of 0.1 second exposures or, if that is not possible, *"custom pixel masks to accurately perform photometry on stars as much as 6 magnitudes brighter than the saturation level"*. This may require User-Generated Processing.

- **Very short or very long exposures –** Special Programs that use very short or very long (>150 sec) exposures could be difficult to calibrate due to having too few (or too few unsaturated) stars.
- **Twilight images –** For Special Programs that obtain images with very bright sky backgrounds (twilight images), it is currently unclear whether they can be processed with the LSST Science Pipelines; User-Generated Processing might be needed (e.g., Sheppard et al. 13).
- **Streaked images –** The full reduction and calibration of data from any Special Programs that use non-sidereal tracking, which produce images with star streaks, is currently beyond the scope of the LSST Science Pipelines; User-Generated Processing would be needed.

4.2 Standard (non-special) processing

Recall that the term "standard processing" is used only within this document, and refers to the image processing pipelines that are designed for, and will be applied to, the Main Survey's observations (Section 2).

Decisions about when to apply standard processing to Special Programs data, or when to include it in the data products for the Main Survey (e.g., if it improves the all-sky coadds), are ultimately left to the discretion of the Rubin Observatory's Data Management and System Performance departments in Rubin Operations.

4.2.1 Prompt Processing and Alert Production

As described in Section 3.3, all visits that *can* be processed by the Prompt pipelines and generate alerts *should* be, in support of time domain and Solar System science goals.

The meaning of "can be processed". This is ultimately left to the discretion of Rubin Data Management in Rubin Operations, but is expected to include all standard and alternative visits in sky regions for which a template image exists. This might also include some non-standard visits (shorter or longer exposures), as long as they can be processed by the Prompt pipeline and an appropriate template image exists. Visits with very short or very long exposure times (or very bright sky backgrounds) might be excluded if they would need specialized algorithms for, e.g., instrument signature removal, difference-imaging, template-generation (Section 4.1). Alert latency requirements do apply to Special Programs data (Section 5.3), and if they cannot be met then Prompt processing is "not possible".

The use of specialized (alternative) template images. If a Special Program's primary science goal requires specialized templates and Prompt processing, the Data Management System will have the capability to load and use an alternative template for some sky regions, based on the image metadata (i.e., the labels described in Section 3.1). However, there would not be enough memory to hold alternative templates for the whole sky.

How Main Survey and Special Program data would co-exist. No "unique and separate" data products for the Special Progams would be produced by standard Prompt processing. Special Programs data that is processed by the Prompt pipeline would contribute to the Prompt data products for the Main Survey as described in Section 3 of the DPDD [LSE-163]. These data products are the results of Difference Image Analysis (DIA), such as the difference images, catalogs of sources detected in difference images (`DiaSources`) and associated static-sky `DiaSources` into `DiaObjects`, and alert packets. Including visits from Special Programs in standard Prompt processing alongside vists from the Main Survey is not, in general, anticipated to affect Main Survey or WFD-specific science goals. For example, analyses for a WFD-only subset could still be done using the observing mode and region labels described in Section 3.1, which would be propogated to difference images, difference-image catalogss, and alerts.

Potential issues with Prompt processing for untiled sequences from Special Programs. There are two potential issues with Prompt processing for DDFs, or any mini-survey that obtains a sequence of untiled images (images at the same pointing or which overlap). For example, a DDF which obtains a hour-long series of about a hundred images at the same coordinates, every few nights for a few months.

1. **DIA Object histories may become too large for the sizing model.** Alert packets contain the full records of all associated `DiaSources` from the past 12 months, but the alert stream bandwidth is sized for the expected histories for Main Survey fields. The Prompt pipeline resources are also sized for the Main Survey, and it might not be possible to load up thousands of epochs at a time. The Data Management team will have to test the realized capabilities of Prompt processing and alert distribtuion, and potentially impose a mitigation strategy such as limiting histories to the last N observations instead of the last 12 months in heavily-observed regions. Alerts from Special Programs data are

subject to the same latency requirements as Main Survey alerts (Section 5.3).

2. **The association of new DIA Sources into DIA Objects may be compromised.** For consecutive images, processing for the second image begins when the processing for the first image is only halfway done. At this point, the `DiaObject` catalog has not yet been updated with the new `DiaSources` detected in the first image. Thus, the `DiaSources` from images one and two for a new transient would not be associated with a single `DiaObject`, but instead would each instantiate a new `DiaObject`.

These two potential issues pose challenges, but are not necessarily showstoppers in processing Special Programs data with the Prompt pipelines. The overall impact on time-domain science would still be positive, even if mitigations are needed for these issues. For example, brokers and users would be able to use the region and observing mode labels in the data as context (i.e., as flags) and avoid including limited-history or potentially-compromised `DiaSources` in their analyses if necessary.

4.2.2 Solar System Processing

Since Solar System Processing takes `DiaSources` as input, any Special Programs images that are processed by the Prompt pipeline could be incorporated into Solar System Processing.

4.2.3 Data Release Processing

Time-domain DIA data products – This includes the results of the annual reprocessing of Main Survey data with Difference Image Analysis (DIA), and the production of Data Release versions of the processed images (single-visit and difference images) and associated catalogs (`DiaSource`, `DiaObject`, `Source`, `ForcedSource`, `DiaForcedSource`, and so on). These data products will be used primarily for time-domain science. They should include Special Programs data for the same reasons as provided in Section 3.3, and with the same considerations as discussed in Section 4.2.1. This processing should use the same template image for a given field.

Static-sky data products – This includes the tessellation and coaddition of Main Survey images and the associated multi-band `Object` catalog and survey property maps. Whether and how to include any Special Programs data in these data products is left entirely to the discretion of the Rubin Data Management team in Operations. As an example, perhaps Special

Programs images will only be included when they assist with uniformity or suppress edge effects or low-order modes in the all-sky photometric solutions.

4.3 Special Processing

As described in Section 3.2, Special Processing should be done by Rubin Data Management to produce unique and separate data products for Special Programs when it is both possible and necessary. This is a requirement (Section 5.2).

In short, *possible* means that original or reconfigured versions of the LSST Science Pipelines can be used , and *necessary* means the primary science goal for Special Program could not be met without the data products produced by the Special Processing.

Joinable tables – Any tables for the unique and separate data products should be joinable to the data products for the Main Survey, when possible. This is a requirement (Section 5.2).

Survey property maps – As for the tesslated all-sky coadded images made from Main Survey data, survey property maps should be made individually as part of any Special Programs that generates tesseladed coadded images. This is *not* a requirement (Section 5.2).

Special Processing timescales – The timescales for Special Processing should be adopted that best serve the primary science goal for the Special Program. For example, nighly-coadd difference-image analysis for DDFs is the most scientifically useful if done on a daily cadence, but the deeply coadded images for DDF fields could be released annually.

4.3.1 Scenarios to illustrate "possible and necessary"

The interpretations of possible and necessary, and the scope of Special Processing, are ultimately left to the discretion of Rubin Data Management in Operations.

Specific examples are provided only to illustrate this interpretation of possible and necessary. These examples do not place limits on what Special Processing might be done.

- **Possible and necessary:** in order to detect high-redshift (faint) galaxies in the DDFs, Rubin Data Management uses the LSST Science Pipelines to deeply coadd images for each field, and store the results of source detection and characterization in unique and separate tables that are included in the annual data release.

- **Possible but not necessary:** a time-domain mini-survey that uses standard visits *could* have separate difference-image analysis object and source catalogs generated, but this is not necessary as the science goals for the mini-survey can be met by processing its data with standard Prompt Processing, and ensuring the data is properly labeled.
- **Possible but not necessary (secondary science goals):** a time-domain mini-survey (or DDF) has a secondary science goal of detecting precursor outbursts for transients, which requires coadding images within windows of days, weeks, and months to reach various depths. This set of custom coadds may be considered as "not necessary" and requiring User-Generated Processing. Note that for the Main Survey data, such custom coadds are also considered as beyond scope and in need of User-Generated Processing.
- **Necessary but not possible:** in order to find the most distant, faint Kuiper Belt Objects in the DDF, a specialized, computationally intensive form of "shift-and-stack" processing is required for detection, but such algorithms are not used by the LSST Science Pipelines and so User-Generated Processing will be needed. A second example is a twilight survey that uses non-standard visits which are outside the boundaries of what the LSST Science Pipelines can process.

Further examples of potential Special Processing for anticipated Special Programs are provided below.

4.3.2 Deep Drilling Fields (DDFs)

As the DDFs will likely be observed with standard or alternative standard visits, Data Management will be able to reconfigure existing pipelines for Special Processing to produce unique and separate DDF data products.

For example, Special Processing for the DDF data products might include:

- nightly-coadded images (24 h)
- nightly-coadded difference images (24 h)

- DiaSource- and DiaObject-like catalogs for the nightly-coadds (24 h)
- deeply-coadded images (all images to date; yearly)
- templates for the nightly-coadded difference images (yearly)
- Source- and Object-like catalogs for the nightly-coadded and deeply-coadded images (yearly)

4.3.3 Short-Exposure Mini-Surveys

As described in [PSTN-055], there are a few short-exposure mini-surveys are under consideration. Two examples are a short exposure map of the sky in $ugrizy$ for calibration, and a Near-Earth Objects (NEO) twilight survey.

Special Processing for short-exposure mini-surveys remains to be determined. Since it falls under the remit of Data Management to perform proper calibration, an evaluatation of whether short exposures for calibration are necessary will be done. If Data Management does find that short-exposure and/or high-sky brightness images can be processed with reconfigured versions of the LSST Science Pipelines, then unique and separate data products could be generated with Special Processing. For the NEO twilight survey, these data products would likely be similar to the Prompt or DIA data products. For the calibration survey, these data products would like be similar to the annual data release tessellated coadds and associated catalog.

4.3.4 Standard Visit Mini-Surveys

Consider a Special Program in which a special region of sky is observed with standard visits but a a special strategy or cadence which is significantly distinct from the Main Survey observing modes, and lasts for a limited amount of time. For example, a short-term survey of the Magellanic Clouds, with dual primary science goals in time-domain and static-sky science.

In cases like this, a set of unique and separate data products with the same formats as the time-domain DIA and static-sky data products described in Section 4.2.3 should be created with Special Processing. They might be released with an annual data release or on an intermediate timescale, e.g., within six months of the conclusion of the mini-survey observations.

4.3.5 Target-of-Opportunity (TOO) Observations

Options for Data Management to process TOO observations, especially during the first year of Operations when the template coverage will be low, are discussed in more detail in RTN-008.

4.4 User-Generated Processing

Science goals that require data products which are not possible to create with the original or reconfigured versions of the LSST Science Pipelines, and/or for which new algorithmic development or significant computational resources are needed, will require user processing and user-generated data products. As described above, custom coadds (e.g., weekly, monthly) are also left to users to generate, as required by their specific science goals.

Computational resources – Users will have access to the LSST Science Pipelines and data processing infrastructure, as well as dedicated computational resources next-to-the-data, via the Rubin Science Platform; LSE-319. Details of the planned "User Batch" facility for data processing are described in DMTN-202. Very computationally intense user processing might require external resources.

Adopting user code or data products – It is expected that some User-Generated pipelines and data products might be "adopted" or "federated" into the LSST Science Pipelines and the Prompt and Data Release data products. Details regarding this are to be provided elsewhere.

Alert production is restricted – User-Generated Processing will not be able to release alert packets in the LSST alert stream. As the latency on processed visit image availability has an 80-hour embargo, no user-generated pipeline will be able to process Special Programs data on a timescale similar to prompt processing and alert production (60 seconds to 24 hours). Thus, no User-Generated Processing may contribute alerts to the LSST alert stream on any timescale.

Further examples of potential User-Generated Processing for anticipated Special Programs are provided below.

4.4.1 Deep Drilling Fields

User-Generated Processing and data products might include, for example, DDF images coadded on custom timescales (e.g., weekly, monthly), or coadded using algorithms outside of the LSST Science Pipelines.

4.4.2 Short-Exposure Mini-Surveys

Short-exposure images obtained during twilight, which will have a very bright sky background unlike other LSST images, might require specialized algorithms to subtract the high sky background which are not available in the LSST Science Pipelines, and might require User-Generated Processing.

Short-exposure images obtained during the night might have too-few stars to satisfy the astrometric and photometric calibration routines in the LSST Science Pipelines, and might require User-Generated Processing.

4.4.3 Mini-Surveys

Mini-surveys with time-domain science goals that aren't met by the Prompt pipelines, e.g., those that require difference imaging with coadded images on an intermediate timescale (e.g., a weekly stack), would require User-Generated Processing.

5 Requirements Related to Special Programs

Appendix D provides a full description of the requirements on the Data Management System related to Special Programs, which are summarized here.

5.1 Metadata

Program metadata that is sufficient to trigger Special Processing should be stored. – In order to support Special Programs processing, the LSST system is required to store metadata that includes program information for every raw image, such as identifiers for images obtained as part of the Main Survey or a Special Program (DMS-REQ-0068). It is required that this metadata be sufficient for Special Programs to trigger their own real-time data processing recipes "whenever possible" (DMS-REQ-0320), and be included in alert packets (DMS-REQ-0274).

5.2 Data Products

Produce unique and special (and joinable) data products when possible. – Rubin Observatory and the LSST system (the observatory and the data management systems) are required to process Special Programs data to produce unique and separate data products "whenever possible" (LSR-REQ-0121). It is a requirement that these Special Programs data products be distinct and joinable with the Prompt and/or Data Release data products (DMS-REQ-0322).

The term "whenever possible" includes cases where the original or reconfigured versions of the LSST Science Pipelines can be run, and excludes cases where the development of new algorithms or the allocation of significant additional computational resources are required (LSR-REQ-0121).

The statement "to produce unique and separate data products" typically refers to producing the same kinds of data products as will be generated by the Prompt and Data Release pipelines (processed visit images, coadded images, difference images, and catalogs of sources and objects for those images).

The term "joinable" means the Special Programs data products can be federated or cross-matched with the relevant Prompt and Data Release data products, and that a column of cross-matched object identifiers is provided to enable table joins.

The size of Special Programs data products should be about 10% of the total. – It is a requirement that the cumulative size of the Special Programs data products generated by Rubin Observatory be no more than ∼10% the size of the Data Release data products (LSR-REQ-0121).

The spirit of this requirement on data volume is that the size be proportional to the fraction of survey time spent on Special Programs.

The derivation of value-added data products, such as HiPS or MOC maps, for Special Programs remains an open question (DMS-REQ-0379, 0383), and is not required. However, they are suggested to be produced with Special Processing where appropriate (Section 4.3).

5.3 Processing

Latency requirements apply to alerts from Special Programs images. – It is a requirement that any Special Programs processing done with the Prompt pipeline (or a reconfigured version of it) is subject to the same timescales and latency constraints of 24 hours for the release of Prompt data products and 1 minute for the transmission of Alert packets (DMS-REQ-0344).

Intermediate timescales should be used to enable science when needed. – It is also a requirement that Special Programs processing be done on timescales intermediate to the Prompt and Data Release processing, "whenever possible" and whenever necessary to enable the intended science goals of the Special Program (LSR-REQ-0032).

Exposure times of 1 second should be processable. – It is a requirement that the LSST system be able to process non-standard visits with short exposure times as low as 1 second, with a discussion note that such short exposures might have degraded image quality (LSR-REQ-0111).

It is not a requirement, but processing for Special Programs by Rubin Observatory is expected to use no more than ∼10% of computational and storage capacity of the Rubin data processing cluster (i.e., proportional to the fraction of survey time spent; Section 6 of LSE-163).

The 10% of the total data processing capacity that Rubin Observatory is required to reserve for *all* User-Generated Processing includes that applied by users to Special Programs data. There is no additional capacity *for users* that will be reserved only for Special Programs data (LSR-REQ-0041).

References

[1] **[Document-11013]**, Becker, A., et al., 2011, *Science White Paper for LSST Deep-Drilling Field Observations Opportunities for Solar System Science*, Document-11013, URL https://ls.st/Document-11013

[2] **[RTN-008]**, Bellm, E.C., 2022, *Rubin Observatory Processing of Gravitational Wave TOO Data*

in the Early Operations Era, RTN-008, URL https://rtn-008.lsst.io/,
Vera C. Rubin Observatory Technical Note

[3] **[RDO-013]**, Blum, R., the Rubin Operations Team, 2020, *Vera C. Rubin Observatory Data Policy*, RDO-013, URL https://ls.st/RDO-013

[4] **[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

[5] **[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

[6] **[DMTN-202]**, Dubois-Felsmann, G., 2021, *Use cases and science requirements on a user batch facility*, DMTN-202, URL https://dmtn-202.lsst.io/,
Vera C. Rubin Observatory Data Management Technical Note

[7] **[LSE-61]**, Dubois-Felsmann, G., Jenness, T., 2019, *Data Management System (DMS) Requirements*, LSE-61, URL https://lse-61.lsst.io/,
Vera C. Rubin Observatory

[8] **[Document-28382]**, Željko Ivezić, Jones, L., Ribeiro, T., 2018, *Call for White Papers on LSST Cadence Optimization*, Document-28382, URL https://ls.st/Document-28382

[9] **[LPM-17]**, Ivezić, Ž., The LSST Science Collaboration, 2018, *LSST Science Requirements Document*, LPM-17, URL https://ls.st/LPM-17

[10] **[LSE-319]**, Jurić, M., Ciardi, D., Dubois-Felsmann, G., Guy, L., 2019, *LSST Science Platform Vision Document*, LSE-319, URL https://lse-319.lsst.io/,
Vera C. Rubin Observatory

[11] **[LSE-163]**, Jurić, M., Axelrod, T., Becker, A., et al., 2023, *Data Products Definition Document*, LSE-163, URL https://lse-163.lsst.io/,
Vera C. Rubin Observatory

[12] LSST Science Collaboration, Marshall, P., Anguita, T., et al., 2017, ArXiv e-prints (arXiv:1708.04058), ADS Link

[13] Sheppard, S.S., Tholen, D.J., Pokorný, P., et al., 2022, AJ, 164, 168 (arXiv:2209.06245), doi:10.3847/1538-3881/ac8cff, ADS Link

[14] **[LDM-151]**, Swinbank, J., Axelrod, T., Becker, A., et al., 2020, *Data Management Science Pipelines Design*, LDM-151, URL https://ldm-151.lsst.io/,
Vera C. Rubin Observatory Data Management Controlled Document

[15] **[PSTN-055]**, The Rubin Observatory Survey Cadence Optimization Committee, 2023, *Survey Cadence Optimization Committee's Phase 2 Recommendations*, PSTN-055, URL `https://pstn-055.lsst.io/`,
Vera C. Rubin Observatory Project Science Technical Note

A Hypothetical Scenarios

Hypothetical examples of Special Programs and the standard, Special, and User-Generated processing to illustrate what *might be done*.

The details of the data acquisition and processing mentioned below are *just illustrative examples* of decisions that have yet to be made.

The steps used to describe the hypothetical processing for each case scenario are:
Step 1. Data acquisition.
Step 2. Standard Prompt processing and Alert Production.
Step 3. Special Processing with reconfigured pipelines.
Step 4. Standard processing for inclusion in Main Survey data products.
Step 5. User-Generated Processing.

A.1 Outer solar system mini-survey

This hypothetical Special Programs processing summary is based on the Becker et al. (2011) white paper to find outer solar system objects with shift-and stack (SAS) [Document-11013].

Step 1. Data acquisition.
In a single night, the 9 adjacent fields in a 3x3 grid are observed with 336×15 second r or g-band exposures (168 standard visits). These observations are repeated 2-3 nights later, and then this 2-night sequence is repeated 3 more times: 1.5 months, 3 months, and 13.5 months later. They are not all at the same RA, Dec, but at selected ecliptic coordinates.

Step 2. Standard Prompt processing and Alert Production.
Each 2×15 second standard visit is processed by the Prompt pipeline and alerts are released

within 60 seconds. Within 24 hours, the `DiaSource` and `DiaObject` catalogs are updated to include the results of Prompt processing of these visits. After 80 hours, the processed visit images and difference images become available. All images and sources originating from this Special Program have region and observing mode labels, e.g., “SP-OSSO”.

The results of Prompt processing are not very relevant for this Special Program’s primary science goal, which requires a year of dispersed observations before the processing pipelines for shift-and-stack can be run. However, including these data in Prompt processing means that they can contribute to LSST’s other time-domain and Solar System science goals.

Step 3. Special Processing with reconfigured pipelines.
None possible. Shift-and-stack processing is beyond the scope of existing algorithms in the LSST Science Pipelines.

Step 4. Standard processing for inclusion in Main Survey data products.
Every year, each 2×15 second standard visit is reprocessed by the DIA data release pipelines and the results are included alongside Main Survey data in the relevant DIA data products (e.g., processed visit images, difference images, associated source catalogs). In the first year after the Special Program is executed, Rubin Data Management finds that 10% of the standard visits from this Special Program had coordinates and image quality that help improve uniformity of the all-sky coadd, and so they are included. In later years, this fraction decreases (remember, this is *hypothetical*). In all data releases, any and all processed images and catalog sources that originate in visits from this Special Program have the same region and observing mode labels, e.g., “SP-OSSO”.

Step 5. User-Generated Processing.
The User-Generated Processing pipeline running the shift-and-stack processing is be set up and submitted for batch processing by the user through the Science Platform or on an external system. The pipeline’s inputs are the processed visit images (and/or difference images) from Prompt processing. User-generated custom algorithms then shift-and-stack the images, and then the LSST Science Pipelines tasks are used to do source detection and characterization and create catalogs. User-generated custom code derives orbital parameters for the detections, and stores them in a user-generated catalog with a similar format to `SSObjects`.

A.2 Deep Drilling Field

Step 1. Data acquisition.
In the COSMOS DDF, the scheduler obtains 10 standard visits in a row in each of the $griz$ filters with a small dither pattern between visits. This happens every other night for a three month season for four years.

Step 2. Standard Prompt processing and Alert Production.
Same as above, but the observing mode label might be, e.g., "SP-DDF-COSMOS".

Step 3. Special Processing with reconfigured pipelines.
First, a template image of appropriate depth for "nightly" difference imaging is created. At the end of each nightly sequence of observations, a pipeline based on recongfigured components of the LSST Science Pipelines is automatically triggered. This pipeline creates nightly coadds in each filter and runs DIA using the template. Alerts are *not* produced, but unique and separate catalogs with the same format as `DiaObject` and `DiaSource` are updated within 24 hours (not a requirement). At the end of each season, deeply coadded images that include all of the DDF's visits from all years are re-generated, along with a separate `Object`-like catalog. All images and catalogs are stored in separate butler collections and TAP tables from the Main Survey data products.

Step 4. Standard processing for inclusion in Main Survey data products.
Every year, each standard visit is reprocessed by the DIA data release pipelines and the results are included alongside Main Survey data in the relevant DIA data products (e.g., processed visit images, difference images, associated source catalogs). Due to their small dither and lack of rotation, not even a single DDF image is used to supplement the Main Survey's all-sky coadd.

Step 5. User-Generated Processing.
In order to achieve a secondary science goal of finding very high-z faint supernovae, a team of users reconfigure the LSST Science Pipelines to create weekly deep coadds of the COSMOS field an appropriate-depth template image, and to run DIA at the end of the season. These data products are stored in separate catalogs with the same format and schema as the `DiaSource` and `DiaObject` tables that are private to the team.

A.3 Short-exposure twilight survey

Twilight observations obtained at, e.g., 60 degrees from the Sun, are particularly well-suited for finding Near-Earth Objects (NEOs).

Step 1. Data acquisition.
At a specified time (or e.g., 6 degree twilight), the scheduler begins a dither pattern of 2-second exposures. Coordinates and exposure times are set by the Sun distance, sky brightness, and desired saturation limits.

Step 2. Standard Prompt processing and Alert Production.
Pending studies of DIA and Alert Production pipeline capabilities to process short-exposure, high sky-background images (see Section 4.1.2).

Step 3. Special Processing with reconfigured pipelines.
Pending studies of the LSST Science Pipelines capabilities to process short-exposure, high sky-background images (see Section 4.1.2).

Step 4. Standard processing for inclusion in Main Survey data products.
These short-exposure, high sky background images would not contribute to the data products created for the Main Survey.

Step 5. User-Generated Processing.
If short-exposure images cannot be processed with the existing DM algorithms, user-generated processing would be needed to reduce the raw data, and to futher detect and characterize sources in the processed images.

B Potential Hardware Boundaries

The potential boundaries on the diversity of data products that could be imposed by limitations from the Rubin Observatory hardware (camera, telescope, and/or site) are considered.

B.1 Filter Changes

The maximum time for filter change is 120 seconds: 30 seconds for the telescope to reorient the camera to its nominal zero angle position on the rotator, and 90 seconds to the camera subsystem for executing the change (OSS-REQ-0293; LSE-30).

Assuming that most Special Programs would be designed to keep overheads <100% and would be using standard 30 second visits, the filter change time indicates that it is likely that at least 4 exposures in a given filter would be obtained between filter changes, but this is not actually a hardware boundary.

The filter change mechanism is designed to undergo a total of 100000 changes over its lifetime, and each filter is designed to support up to 30000 changes over its lifetime, where lifetime is 15 years.

That is an average of ~27 changes per day, some of which would occur in they day during calibrations (estimate, ~10) and the rest at night.

As stated in the filter change memorandum (`ls.st/spt-494`), *"the system could support as many changes involving the 5 filters loaded in the carousel as desired, without any practical limitation"*.

B.2 Filter Carousel Loads

The filter carousel can hold five of the six LSST filters at a time. The system is designed to support 3000 loads in 15 years (`ls.st/spt-494`). Filter loads are only done in the day, and there will never be data in more than five filters in a given night.

B.3 Exposure Times

The minimum exposure time is 1 second, with a stretch goal of 0.1 seconds (OSS-REQ-0291; LSE-30). The maximum exposure time is not restricted.

B.4 Readout Time

The readout time is 2 seconds, and would be significant overhead on short exposures.

B.5 Inter-Image Time

Images with exposure times < 15 seconds *might* still have to be separated by 15 seconds for thermal tolerance; i.e., the minimum readout rate might be one image every 15 seconds, regardless of exposure time (OSS-REQ-0291; LSE-30).

As discussed in Jira ticket DM-12573, the main issue is thermal and is related to the shutter, both the motors and the brakes; an elevated Camera skin temperature would affect image quality.

As of 2022, early tests suggest that a sustained (30 minutes) sequence that increases the heat load by large factors would not work, but further functional testing of the system once the Camera was fully assembled are needed for full characterization of the issue.

This potential 15 interval between images is also a potential hardware boundary on the potential diversity of data products.

B.6 Telescope Slew

As described in Document-28382, large slews would have considerable overheads, but there are no hardware boundaries on the size of a single slew or the accrued slew distance.

B.7 Pointing and Sidereal Tracking

The specifications for the telescope's pointing and tracking in LSE-30 indicate that <0.2 arcsecond precision in field pointing (OSS-REQ-0302) and <1 arcsecond in open-loop tracking (OSS-REQ-0303) would not be possible, but guiding would improve the latter (OSS-REQ-0305).

Furthermore, obtaining the *exact same* alignment of the pixel grid in RA-Dec *"would put demands on the camera rotator that were not planned"*[4].

[4] As per C. Claver's comments in ticket DM-12573.

B.8 Non-Sidereal Tracking

The requirement that the LSST system be able to perform non-sidereal tracking is set by OSS-REQ-0380 in LSE-30. This capability will include angular rates of up to 220 arcseconds per second in both azimuth and elevation.

B.9 Camera Rotation

The requirements on the rotator's capabilities do not set any limits on the per-night or total lifetime rotation (OSS-REQ-0301, -0300; LSE-30) which might put boundaries on the distance between successive visits or the ability to jump between two widely separated fields.

Currently, there are no hardware boundaries imposed by camera rotation constraints on the potential diversity of data products.

C Potential Processing Boundaries

The capability of the LSST Science Pipelines to process diverse data is explored below.

Note that processing boundaries might ultimately be defined not by what is technically possible, but by the resulting image quality parameters, e.g., the number of stars with sufficient flux for photometric calibration.

Furthermore, the processing boundaries might not be fully constrained until the final performance of the LSST Science Pipelines, as described in the Data Management Applications Design, LDM-151) document, is fully characterized.

Summary.

- **Short or long exposures.** Very short (<2 sec) exposures could be difficult to process due to an incompletely-formed PSF. Short (or long, >150 sec) exposures could be difficult to calibrate due to having too few (or too few unsaturated) stars.
- **Bright backgrounds (twilight).** It is currently unclear whether images with very bright

sky backgrounds (twilight images) can be processed with the LSST Science Pipelines, or whether user generated pipelines will be needed.

- **Non-sidereal tracking.** The full reduction and calibration of images obtained with non-sidereal tracking, in which the stars are streaked, is currently beyond the scope of the LSST Science Pipelines, and will require a user generated pipeline.

C.1 Exposure Times

Images which deviate significantly from the 15 second duration for the Main Survey may encounter issues in the instrument signature removal routine, in the correction for differential chromatic refraction, in the difference imaging analysis pipeline, and/or in the photometric and astrometric calibrations due to a differently sampled set of standard stars per CCD.

C.1.1 Short Exposures (Non-Standard Visits of $<$30 sec)

The LSST System Requirements document states that *"The LSST shall be capable of obtaining and processing exposures not taken in a standard visit mode including those with a minimum exposure time of* `minExpTime`", which is 1 second (stretch goal 0.1 seconds; LSR-REQ-0111 in LSE-29).

However, for exposure times there are other considerations, as changing the exposure time also affects the photometric and astrometric calibrations. Assuming that 1 second exposure can be reduced and calibrated, its detected point sources will span a dynamic range of $r\approx$ 13 to 21 magnitudes. A template image built on 15 second exposures will saturate at $r\approx$15.8 mag, but this still leaves stars between 15.8 and 21.0 magnitudes to be used in the PSF-matching (and all other filters have a similarly large overlap).

In order for an image to be successfully PSF-matched to the template, the PSF must be well formed (no speckle pattern), and have a spatial variance that the pipeline is capable of modeling (be smoothly varying on some minimal scale). As a simple demonstration, Figure 1 shows that perhaps exposure times shorter than 2 seconds do not have a well-formed PSF (using the centroid of a 2D Gaussian fit as a proxy for "well-formed").

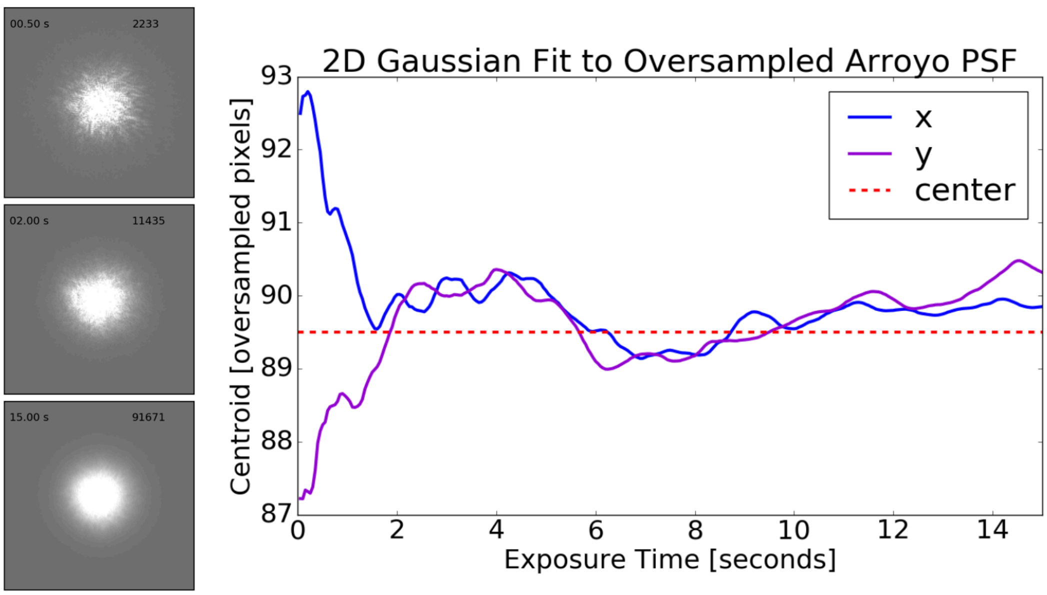


FIGURE 1: At left, Arroyo atmosphere-only simulated PSF for LSST (with oversampled pixels) with exposure times of 0.5, 2, and 15 seconds (top to bottom), courtesy of Bo Xin. At right, blue and purple lines show the location of the centroid derived from a 2D Gaussian fit to the PSF as a function of exposure time, with the red dashed line showing the true center. We can see that for exposure times greater than 2 seconds, the centroid converges near its true value.

C.1.2 Long Exposures (Non-Standard Visits of >30 sec)

There is no maximum exposure time specified for an LSST image. Given that the template image will be a stack of at least a year or two of data, processing a 5–10 times deeper single image through the difference imaging pipeline should be fine.

However, a 2×150 second exposure would saturate at $r \approx 18.3$, perhaps leaving too few stars overlapping with e.g., templates or Main Survey images, for astrometric and photometric calibrations.

Furthermore, cosmic-ray rejection completeness might be reduced for longer exposures (unknown), which could impact the quality of a difference image and the detected sources. Additionally, any system qualities that vary on short (but > 30 second) timescales could inhibit photometric calibration (e.g., tracking).

C.2 Twilight Images with a Bright Background

Images obtained during twilight for scientific purposes are also likely to have shorter exposure times, and so the issues described in Section C.1 also apply here.

Whether or not bright-background images can (or shall) be fully processed – reduced, calibrated, background-subtracted, and delivered with astrometric and photometric solutions – or whether this will require a user generated pipeline, remains to be determined (see also the example in Section A.3). This may depend on the exposure time and the number of stars available in the image.

C.3 Images Obtained with Non-Sidereal Tracking

Non-sidereal tracking leads to images in which stars are streaked, but the moving object appears as a point source.

Full processing – providing reduced, calibrated, background-subtracted images that are delivered with astrometric and photometric solutions – of these images is beyond the scope of the DM pipelines as it would require the development of new algorithms, and will need to be done as a user generated pipeline. The first steps of such a pipeline, such as Instrument Signature Removal, will probably be possible to achieve by reconfiguring the relevant DM software tasks.

D Documentation Review for Requirements Related to Special Programs

This review is limited to requirements related to Special Programs processing or data products, or requirements that constrain the diversity of images (e.g., exposure time limits).

Updates to Rubin documents related to Special Programs that were motivated by past versions of this DMTN were made via LCR-1309 and LCR-2265.

D.1 Science Requirements Document (SRD)

Version 5.2.4 (revision 2018-01-30), [LPM-17].

Section 3.4 "The Full Survey Specifications" states the SRD's assumption that 90% of the total available survey time would be spent on the main survey, and that the remaining 10% would be spent *"to obtain improved coverage of parameter space ... [or to] observe special regions"*.

D.2 LSST System Requirements (LSR)

Version 7.1 (revision 2020-03-05), LSE-29.

Note that Version 5 (2018-06-26) was an update for LCR-1309, which added requirements, specifications, and discussions regarding the processing of Special Programs data based on earlier versions of this DMTN.

In Section 1.5.1.3, "Processing Data from Special Programs", LSR-REQ-0122 is a requirement LSR-REQ-0122
that the LSST system *"shall deliver unique and separate data products for visits from Special Programs"* whenever possible, and that they *"shall be delivered on timescales intermediate"* to the Prompt and Data Release timescales *"when this enables the intended science of the Special Program"*. The discussion clarifies that *"the term 'whenever possible' includes cases where the Data Management System can run original or reconfigured versions of existing pipelines, and excludes*

cases where the development of new algorithms, or the allocation of significant additional computational resources, are required".

In Section 2.4.1.1.2, "Non-Standard Visit", LSR-REQ-0111 requires that the LSST system *"be capable of obtaining and processing exposures not taken in a standard visit mode including those with a minimum exposure time of"* 1 second (`minExpTime`). The discussion notes that *"non-standard visit exposures may possibly be degraded in some aspects of performance (e.g. cosmic ray rejection on visits consisting of a single exposure), and might be incompatible with difference imaging and alert production (e.g., short exposures in which the PSF is not fully formed)"*. LSR-REQ-0111 minExpTime

The requirement in Section 1.5.1.3 is echoed in Section 2.6.1.1, "Organization of Data Products", in which LSR-REQ-0032 is a requirement that the data processing system provide the means for three 'classes' of data products on different timescales (Prompt, Data Release, and User-Generated), and also to provide a means for processing Special Programs data because the *"science goals of Special Programs may require that their processed data products be made available in an additional fourth class, and possibly with intermediate timescales"*. LSR-REQ-0032

In Section 2.6.1.1.3, "Level 3 Data Products", LSR-REQ-0041 specifies that the LSST system *"shall support"* User-Generated data products. The discussion clarifies that "*there will be technical limits on DM's ability to meet this requirement, such as cases where an intensive amount of additional computational resources is required, because only 10% of the total computational system is allocated for user processing"*. This level of support applies also to user processing of Special Programs data. See also the reference to LSR-REQ-0055 below. LSR-REQ-0041

Section 2.6.1.1.4, "Data Products for Special Programs", LSR-REQ-0121 specifies that the LSST system *"shall produce unique and separate Data Products as the result of processing data from Special Programs whenever possible, on a timescale that enables the intended science goals of the Special Program. The cumulative size of the online Special Programs data products shall be no more than 10% of the size of the DRP data products from the most recent data release"*. The discussion clarifies that *"the term 'whenever possible' includes cases where the Data Management System can run original or reconfigured versions of existing pipelines, and excludes cases where the development of new algorithms, or the allocation of significant additional computational resources, are required. The cumulative size of the Special Programs data products is capped at 10% of the most recent DR because this matches the expected fractional survey area of Special Programs compared to the main survey"*. LSR-REQ-0121

In Section 2.7.1.6, "Community Computing Services", LSR-REQ-0055 requires that the LSST system *"shall provide and maintain an amount of computing capacity equivalent to at least"* 10% (`userComputingFraction`) *"of the total LSST data processing capacity (computing and storage) for the purpose of scientific analysis of LSST data and the production of"* User-Generated data products. The discussion clarifies that the scope of this service remains to be determined. This level of computational resources includes user processing of Special Programs data. LSR-REQ-0055 userComputingFraction

In Section 3.1.3.1, "Survey Time Allocation", LSR-REQ-0075 requires that the *"survey performance requirements shall be met utilizing approximately 90% of the historically available observing time, leaving the remaining time available for yet to be defined special programs"*. LSR-REQ-0075

D.3 Observatory System Specifications (OSS)

Version 19.3 (revision 2022-12-12), LSE-30.

Note that Version 13 (2018-06-26) was an update for LCR-1309, which added requirements, specifications, and discussions regarding the processing of Special Programs data based on earlier versions of this DMTN.

In Section 2.2.3.1, "Standard Operating States", OSS-REQ-0044 specifies that *"the LSST observatory system shall be designed and constructed to support ... manual observing - used for specific non-scheduler driven observing to support system verification and testing or specialized science programs"*. Although most Special Programs will be executed via the survey scheduler as part of *"fully automated observing"*, manual observing might be necessary for, e.g., target-of-opportunity Special Programs. OSS-REQ-0044

Section 3.1.5.1.2, "Data Products Handling for Special Programs", OSS-REQ-0392 is a flow-down of requirements from the LSR (0122, 0075, and 0121; LSE-29), and specifies that *"the handling of data products from Special Programs shall be compliant with the approach defined in LSE-163"*. OSS-REQ-0392

In Section 3.6.1.3, "Continuous Exposures", OSS-REQ-0319 requires that *"The Observatory shall be capable of continuous operation throughout a night with the interval between successive visits equal to the FPA readout time"*. The discussion clarifies that *"this mode of observing is needed to support observations when the telescope is not being re-pointed. For example observing "deep drilling" fields..."*. OSS-REQ-0319

In Section 3.6.1.4, "Minimum Exposure Time", OSS-REQ-0291 specifies that *"the camera shall be able to obtain a single exposure with an effective minimum exposure time of no more than"* 1 second (`minExpTime`) *"with a goal of an effective minimum exposure time of"* 0.1 seconds (`minExpTimeGoal`). The discussion clarifies that *"if the exposure is shortened from the 15 second nominal, the spacing between successive exposures should be extended to maintain the average readout rate consistent with a 15 second exposure"*, which may increase the overheads of Special Programs using short exposure times. The discussion also clarifies that *"if the exposure is lengthened from the 15 second nominal, the thermal stability may be affected, which may affect photometric accuracy"*. OSS-REQ-0291 minExpTime minExpTimeGoal

In Section 3.6.1.5, "Publish Visit Type", OSS-REQ-0384 specifies that *"the OCS [Observatory Control System] shall configure the [Data Management System] DMS (in particular Prompt Processing) with the type of visits to be processed: Standard, Alternate, or a specific type of Non-Standard"*. The discussion clarifies that this allows the Prompt processing pipeline to be reconfigured on-the-fly in order to incorporate non-standard visits from, e.g., Special Programs. The time required for reconfiguration might introduce some latency or cause some images to not be processed by the Prompt pipeline. OSS-REQ-0384

In Section 3.6.2.1.2, "Maximum time for operational filter change", OSS-REQ-0293 specifies that *"the camera system shall provide the capability of changing the operational filter with any other internal filter in a time less than"* 120 seconds (`tFilterChange`). This would impose a large overhead on, e.g., a Special Program that changes filters often without slewing. See also OSS-REQ-0295, Appendix B of this document, and/or the filter change memorandum (`ls.st/spt-494`), for more information about the total lifetime number of filter changes. OSS-REQ-0293 tFilterChange OSS-REQ-0295

In Section 3.6.3.1, "Absolute Pointing", OSS-REQ-0298 specifies that *"the LSST shall point to a defined set of sky coordinates with an RMS accuracy of"* 2 arcseconds (`absPointErr`). OSS-REQ-0298 absPointErr

In Section 3.6.3.3, "Rotator tracking Time, OSS-REQ-0301 specifies that *"the LSST shall be able to maintain field rotation tracking over a period of at least"* 1 hour (`rotTrackTime`). The discussion clarifies that this *"is driven by the need to conduct extended 'deep drilling' observations on a single field"*. OSS-REQ-0301 rotTrackTime

In Section 3.6.3.5, "Offset Pointing", OSS-REQ-0302 specifies that *"the LSST shall be capable of offset pointing within a single field-of-view with a precision of no more than"* 0.2 arcseconds (`offsetPointingErr`). OSS-REQ-0302 offsetPointingErr

In Section 3.6.3.6, "Open Loop Tracking", OSS-REQ-0303 specifies that *"The LSST shall be capable of open loop tracking without the assistance of real time optical feedback to an accuracy of"* 1.0 arcseconds (`openTrackErr`) *"over any 10 minute duration during normal night time operations"*. Note that the open loop tracking requirement is *without guiding*. OSS-REQ-0303 openTrackErr

In Section 3.6.3.10, "Non-Sidereal Tracking", OSS-REQ-0380 specifies that *"the LSST system shall be capable of tracking in an arbitrary direction on the sky along a parametric RA(t) and DEC(t) trajectory, at angular rates of up to"* 220 arcseconds per second (`nonsiderealAngularRateEl` and `nonsiderealAngularRateAZ`) *"with a tracking error not to exceed"* 0.5 arcseconds per minute (`nonsiderealTrackingError`). The discussion notes that *"this is standard capability for modern telescopes"*, but might be relevant to some Special Programs. OSS-REQ-0380 nonsiderealAngularRateEl nonsiderealAngularRateAZ nonsiderealTrackingError

D.4 Data Management Subsystems Requirements (DMSR)

Version 9 (revision 2021-02-12), LSE-61.

Note that Version 8.3 (2020-05-04) was an update for LCR-2265, which updated requirements, specifications, and discussions regarding the processing of Special Programs data based on earlier versions of this DMTN.

In Section 1.2.3, "Raw Science Image Metadata", DMS-REQ-0068 specifies that *"for each raw science image, the DMS shall store image metadata"* including *"Program metadata (identifier for main survey, deep drilling, etc.)"*. The discussion clarifies that *"the program metadata should be sufficient to associate an image with a specific Special Program so that DMS-REQ-0320 and DMS-REQ-0397 can be satisfied"*. DMS-REQ-0068

In Section 1.3.13, "Alert Content", the discussion for DMS-REQ-0274 explains that the *"program and/or scheduler metadata"* included in an alert packet *"should be sufficient to identify whether the image is associated with a Special Program (such as an in-progress Deep Drilling Field)"*. DMS-REQ-0274

In Section 1.4.18.1, "Produce All-Sky HiPS Map", the discussion for DMS-REQ-0379 raises the point that generating separate HiPS maps for Special Programs (e.g., DDFs) remains an open question. DMS-REQ-0379

In Section 1.4.18.5, "Produce MOC Maps", DMS-REQ-0383 specifies that Data Release processing *"shall include the production of Multi-Order Coverage maps for the survey data"*, and that DMS-REQ-0383 *"additional MOCs SHOULD be produced to represent special-programs datasets"*. It is noted that a separate technical note would be created to define these MOCs.

The bulk of the DMS's requirements related to Special Programs are in Section 1.6 of the DMSR.

In Section 1.6.1, "Processing of Data From Special Programs", DMS-REQ-0320 specifies that *"it shall be possible for special programs to trigger their own data processing recipes, during the night instead of the nightly Alert Processing (but the recipes may still issue Alerts), or on alternative timescales"*. The discussion clarifies that the *"LSST will provide these recipes ... when possible, which includes cases where DM can run original or reconfigured versions of existing pipelines, and excludes cases where the development of new algorithms, or the allocation of significant additional computational resources, are required. An example of an alternative timescale is a nightly trigger to coadd all the deep-drilling field images. Decisions about which recipes are applied to which Special Programs will be made by the Operations team, after consideration of the scientific goals, computational resources, and data rights policy"*. This requirement is derived from OSS-REQ-0392, which is essentially a flow-down of requirements from the LSR (0122, 0075, and 0121). DMS-REQ-0320

In Section 1.6.2, "Prompt/DR Processing of Data from Special Programs", DMS-REQ-0397 specifies that *"it shall be possible for special programs data to be processed with the prompt and/or annual-release pipelines alongside data from the main survey"*. The discussion further clarifies that *"the data from Special Programs should only be included ... when it is (a) possible ... to do so without additional effort and (b) beneficial to the LSST's main science objectives. Decisions about which data are included ... will be made by the Operations team"*. This requirement is also derived from OSS-REQ-0392, which is essentially a flow-down of requirements from the LSR (0122, 0075, and 0121). DMS-REQ-0397

In Section 1.6.3, "Level 1 Processing of Special Programs Data", DMS-REQ-0321 specifies that *"all [Prompt] processing from special programs shall be completed before data arrives from the following night's observations"*. This is essentially adding a quantifier to DMS-REQ-0397, to specify that *"when it is (a) possible ... to do so"* means when it is possible to complete the processing before the next night's observations. This requirement is also derived from OSS-REQ-0392, which is essentially a flow-down of requirements from the LSR (0122, 0075, and 0121). DMS-REQ-0321

In Section 1.6.4, "Constraints on Level 1 Special Program Products Generation", DMS-REQ-0344 specifies that *"the publishing of [Prompt] data products from Special Programs shall be sub-* DMS-REQ-0344 *ject to the same performance requirements of"* 24 hours (`L1PublicT`) for the release of Prompt data products and 1 minute (`OTT1`) for the transmission of Alert packets. This is essentially a more detailed version of DMS-REQ-0321 which includes the Alert production timescale. This requirement is also derived from OSS-REQ-0392, which is essentially a flow-down of requirents from the LSR (0122, 0075, and 0121). L1PublicT OTT1

In Section 1.6.5, "Special Programs Database", DMS-REQ-0322 specifies that *"data products for special programs shall be stored in databases that are distinct from those used to store standard [Prompt] and [Data Release] data products"* and that *"it shall be possible for these databases to be federated ... to allow cross-queries and joins"*. This requirement is also derived from OSS-REQ-0392, which is essentially a flow-down of requirements from the LSR (0122, 0075, and 0121). DMS-REQ-0322

In Section 4.1.16, "Level 2 and Reprocessed Level 1 Catalog Access", DMS-REQ-0313 specifies that *"the DMS shall maintain ... versions of the most recent catalogs generated from Special Programs data"*. As with all LSST data, *"there is no requirement for older data releases to be queryable"*. DMS-REQ-0313

D.5 Data Management Applications Design (DMAD)

Version 4.3 (revision 2020-11-10), LDM-151.

The DMAD is not a requirements document. Instead, it describes the scientific design of the LSST Science Pipelines: the algorithms and software that will be implemented to meet the requirements for processing the LSST data.

Special Programs are only mentioned a few times, either as a potential source of single-snap visits or as a potential source of reference images or catalogs (e.g., training sets).

As described above (e.g., LSR-REQ-0122), the LSST system shall deliver unique and separate data products for visits from Special Programs whenever this (1) enables the intended science of the Special Program and (2) can be accomplished using the original or reconfigured versions of the LSST Science Pipelines. For cases in which the development of new algorithms or the allocation of significant additional computational resources are required to produce Special Programs data products, User-Generated pipelines and processing will be necessary.

The DMAD can be used as the reference document to decide whether a given Special Program will require User-Generated pipelines and processing.

D.6 Data Products Definitions Document (DPDD)

Version 3.6 (revision 2021-12-17), LSE-163.

Section 6 describes the data products for Special Programs. The DPDD is not a requirements document; Section 6 summarizes the requirements presented above and does not introduce any new constraints or new information about Special Programs.