Rubin Observatory

Vera C. Rubin Observatory
Data Management

Characterizing `DIASource` Detection Efficiency

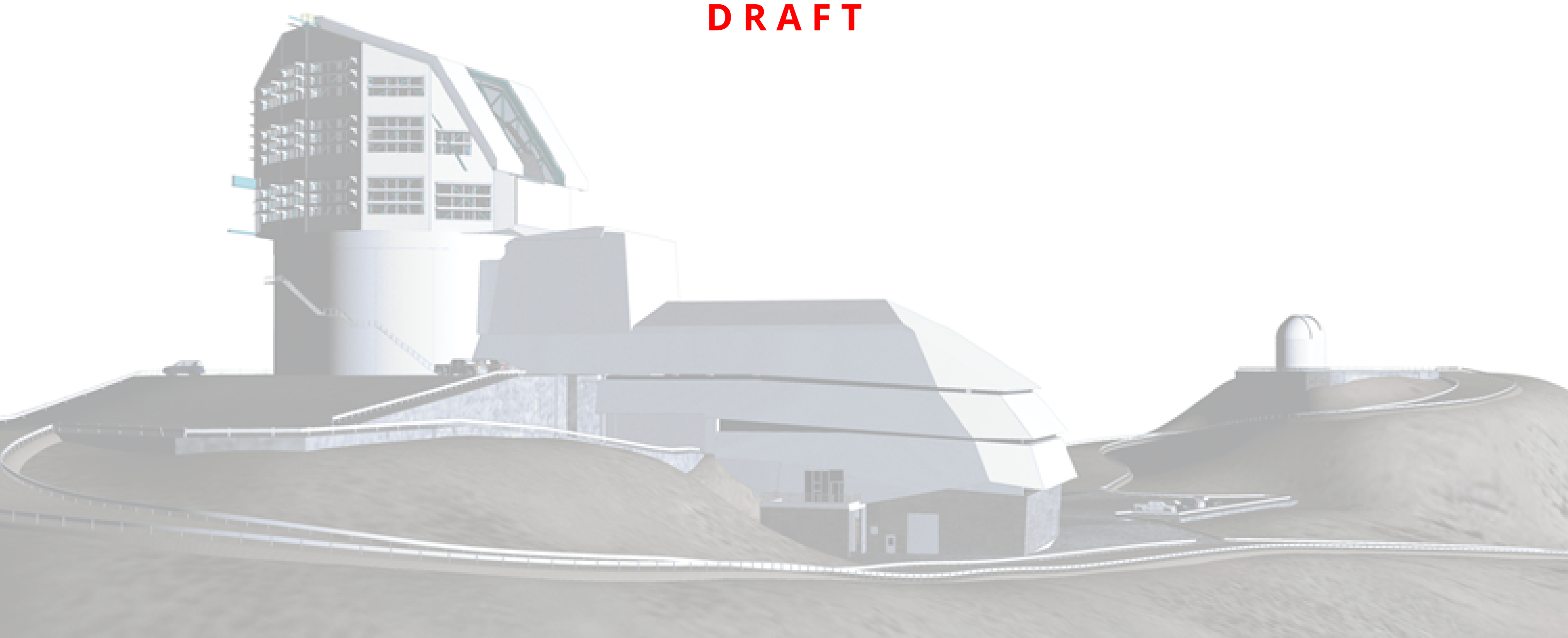
Draft

M. L. Graham et al., and the DM SST

DMTN-TBD

Latest Revision: 2022-04-09

D R A F T



Rubin Observatory

Abstract

The goals of this study are to: (1) validate that the planned DM data products which characterize `DIASource` detectability will meet the scientific needs of the community; (2) ensure that the requirements related to the characterization of `DIASource` detectability are adequately flowed down, and that DM's plans are accurately expressed in the DM documentation; and (3) inform the community about the planned DM data products related to `DIASource` detectability.

Draft

Change Record

Version	Date	Description	Owner name
0.0	2018-11-01	Internal working document.	Melissa Graham
0.1	2019-09-09	Updated to represent DM's plans.	Melissa Graham
1	2022-04-dd	Update and release.	Melissa Graham

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

Draft

Rubin Observatory

Contents

Rubin Observatory

Draft

Rubin Observatory

Characterizing `DIASource` Detection Efficiency

1 Introduction

Any astrophysical question which asks, e.g., *"How often?"* or *"How many?"* about transient phenomena, such as population studies or occurrence rates, needs to know a survey's *detection efficiency*: the probability that a source is detected (or in other words, the fraction of all sources that are detected).

Detection efficiencies are required for many of the core science pillars of the LSST, such as transient phenomena, cosmology, and Solar System studies; many other surveys use populations of synthetic sources injected into the data in order to characterize the detection efficiency (Appendix A). This should be done for the LSST as well – deriving detection efficiencies is too large and complicated a task to leave for the science community.

Rubin Observatory documentation contains no requirements related to producing or serving detection efficiencies for transient `DIASources` in difference images (Appendix B), but does contain requirements about using source injection to characterize detected source *spuriousness* (see also Appendix C).

This document describes how Rubin Observatory should use its synthetic source injection algorithms to measure and provide detection efficiencies for transient point sources detected by the LSST via difference image analysis (DIA).

Readers should refer to the Data Products Definitions Document [LSE-163] for more information about DIA, difference images, and `DIASources`. Detection efficiencies of extended difference-image sources (e.g., light echoes, trailed moving objects), or for variable or static point-sources in direct images, are beyond the scope of this document.

2 Proposed Detection Efficiency Matrix

Detection: as described in the SRD and DPDD, sources in difference images with a signal-to-noise ratio $SNR > transSNR = 5$ will be considered *detected* and will become a `DIASource`.

Detection Efficiency: The probability that a true astrophysical source of a given magnitude is detected and becomes a `DIASource`. In other words, the fraction of all true astrophysical sources of a given magnitude that are detected and become `DIASources`.

Characterizing detection efficiency requires *knowing* which of the detected `DIAsources` are astrophysical, and the total number of astrophysical sources in the field of the image. This would only be possible if "truth" is somehow known from, e.g., co-temporal imaging data of superior quality – but of course that almost never exists because it is inefficient to duplicate data.

Instead, the detection efficiency is typically characterized by simulating and injecting synthetic sources into images, using a point-spread function shape like real astrophysical sources, and then processing the images with the survey's difference image analysis pipeline and measuring the fraction recovered.

The detection efficiency can be characterized as $\eta(m)$, where m is the apparent magnitude of the time-changing component with respect to the template image, and η is a value between 0 and 1 that represents the probability that the source would be detected in the difference image. As described in Appendix A, η depends on more than just m, and is a function of the parameters ($\vec{P}$) listed in Table 1.

An accurate measure of $\eta(m, \vec{P})$ for every pixel of every difference image is technologically unfeasible. Instead, an analytic model for η as a function of both m and parameters $\vec{P}$ (some of which may be correlated) can be built with synthetic source injection, and then applied by users.

This means that the synthetic sources should be simulated over a range of parameters $\vec{P}$, but as described in Appendix A the simulated sources do not need to accurately represent astrophysical transient types, colors, redshifts, durations, light curves, etc. That aspect of the analysis is best left to the user to handle during the MC simulation stage for their particular type of object.

TABLE 1: A description of the image and source parameters ($\vec{P}$) that can affect the detection efficiency (η) of point sources in a difference image (per filter).

Parameter	Description
Surface Brightness	Typically, η decreases for objects embedded in brighter host galaxies.
Static-Source Offset	Sometimes, η decreases for objects that are near (i.e., overlap the point-spread function of) static sources (e.g., stars, galaxy cores, especially if cuspy in profile).
CCD Location	With some instruments, η decreases near the CCD edges due to distortion.
Image FWHM	The value of η can decrease for extreme FWHM differences from the template (i.e., very good or very poor seeing).
Image Airmass	LSST images will experience differential chromatic refraction which affects image subtraction [DMTN-037], and thus potentially also η.
Sky Brightness	Typically, η decreases when the sky background is bright or has a strong gradient (e.g., during twilight, near the moon).
Sky Cloud Cover	Extinction will affect η by degrading the image magnitude limit.

Rubin Observatory

References

[1] Brink, H., Richards, J.W., Poznanski, D., et al., 2013, MNRAS, 435, 1047 (arXiv:1209.3775), doi:10.1093/mnras/stt1306, ADS Link

[2] Choi, Y., Gibson, R.R., Becker, A.C., et al., 2014, ApJ, 782, 37 (arXiv:1312.4957), doi:10.1088/0004-637X/782/1/37, ADS Link

[3] **[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

[4] **[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

[5] Dilday, B., Kessler, R., Frieman, J.A., et al., 2008, ApJ, 682, 262 (arXiv:0801.3297), doi:10.1086/587733, ADS Link

[6] **[LSE-61]**, Dubois-Felsmann, G., Jenness, T., 2018, *LSST Data Management Subsystem Requirements*, LSE-61, URL https://ls.st/LSE-61

[7] Figuera Jaimes, R., Bramich, D.M., Skottfelt, J., et al., 2016, A&A, 588, A128 (arXiv:1512.07913), doi:10.1051/0004-6361/201527641, ADS Link

[8] Frohmaier, C., Sullivan, M., Nugent, P.E., Goldstein, D.A., DeRose, J., 2017, ApJS, 230, 4 (arXiv:1704.02951), doi:10.3847/1538-4365/aa6d70, ADS Link

[9] Frohmaier, C., Sullivan, M., Maguire, K., Nugent, P., 2018, ApJ, 858, 50 (arXiv:1804.03103), doi:10.3847/1538-4357/aabc0b, ADS Link

[10] Frohmaier, C., Sullivan, M., Nugent, P.E., et al., 2019, MNRAS(arXiv:1903.08580), doi:10.1093/mnras/stz807, ADS Link

[11] Graham, M.L., Sand, D.J., Bildfell, C.J., et al., 2012, ApJ, 753, 68 (arXiv:1205.0015), doi:10.1088/0004-637X/753/1/68, ADS Link

[12] Holoien, T.W.S., Kochanek, C.S., Prieto, J.L., et al., 2016, MNRAS, 455, 2918 (arXiv:1507.01598), doi:10.1093/mnras/stv2486, ADS Link

[13] Hung, T., Gezari, S., Cenko, S.B., et al., 2018, ApJS, 238, 15 (arXiv:1712.04936), doi:10.3847/1538-4365/aad8b1, ADS Link

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

[15] **[LSE-163]**, Jurić, M., et al., 2017, *LSST Data Products Definition Document*, LSE-163, URL `https://ls.st/LSE-163`

[16] Kankare, E., Mattila, S., Ryder, S., et al., 2012, ApJ, 744, L19 (`arXiv:1112.0777`), doi:10.1088/2041-8205/744/2/L19, ADS Link

[17] Kessler, R., Marriner, J., Childress, M., et al., 2015, AJ, 150, 172 (`arXiv:1507.05137`), doi:10.1088/0004-6256/150/6/172, ADS Link

[18] Lee, C.H., Riffeser, A., Seitz, S., Bender, R., Koppenhoefer, J., 2015, ApJ, 806, 161 (`arXiv:1504.07246`), doi:10.1088/0004-637X/806/2/161, ADS Link

[19] Pérez-Torres, M.A., Romero-Cañizales, C., Alberdi, A., Polatidis, A., 2009, A&A, 507, L17 (`arXiv:0909.3959`), doi:10.1051/0004-6361/200912964, ADS Link

[20] Sako, M., Bassett, B., Becker, A., et al., 2008, AJ, 135, 348 (`arXiv:0708.2750`), doi:10.1088/0004-6256/135/1/348, ADS Link

[21] Sand, D.J., Graham, M.L., Bildfell, C., et al., 2012, ApJ, 746, 163 (`arXiv:1110.1632`), doi:10.1088/0004-637X/746/2/163, ADS Link

[22] **[DMTN-037]**, Sullivan, I., 2017, *DCR-matched template generation*, DMTN-037, URL `https://dmtn-037.lsst.io`,
LSST Data Management Technical Note

[23] **[LDM-151]**, Swinbank, J.D., et al., 2017, *Data Management Science Pipelines Design*, LDM-151, URL `https://ls.st/LDM-151`

[24] Trevese, D., Boutsia, K., Vagnetti, F., Cappellaro, E., Puccetti, S., 2008, A&A, 488, 73 (`arXiv:0805.2948`), doi:10.1051/0004-6361:200809884, ADS Link

[25] Villforth, C., Koekemoer, A.M., Grogin, N.A., 2010, ApJ, 723, 737 (`arXiv:1008.3384`), doi:10.1088/0004-637X/723/1/737, ADS Link

Draft

A Scientific Examples

A.1 Summary

Detection efficiencies are required for many of the core science pillars of the LSST, as described below. Many other surveys use populations of synthetic sources injected into the data in order to characterize the detection efficiency, as described in the examples provided in A.2.

In consideration of the scientific motivation for detection efficiencies, and what has worked for other surveys, two points are clear.

1. The population of synthetic sources injected in order to characterize the detection efficiency should have similar distributions as real astrophysical phenomena in terms of brightness and location (e.g., proximity to static-sky sources, field crowdedness, surface brightness), and the subset of images used should sample the distributions of, e.g., image quality (seeing), sky brightness, and airmass.
2. The simulated fakes do not need to accurately represent astrophysical transient types, colors, redshifts, durations, light curves, etc., or be planted in sequential images in a correlated way to represent real light curves. That aspect of the analysis is best left to the user to handle during the MC simulation stage for their particular type of object.

A.2 Transients

Transient events such as stellar explosions (supernovae, kilonovae) and tidal disruption events (TDEs; stars destroyed by close passage to a supermassive black hole) occur once and do not repeat. Since most transient progenitors are stars, they are most often found in high surface brightness environments (i.e., galaxies; their spatial distribution “follows the light") and require difference imaging in order to be discovered, and thus detection efficiencies in order to characterize their occurrence rates. For example, transient occurrence rates as a function of environment can constrain their progenitor star characteristics, which requires that detection biases be well known, as does understanding selection biases in transient samples (e.g., when using Type Ia supernova as cosmological standard candles).

The four surveys mentioned below have either inserted fakes into all of their live images (SDSS-II and DES) or into a representative set of images at a later time (CFHT and PTF), to ensure that detection efficiencies can be determined for the full range of image parameters. In all cases, the fakes were simulated with parameter distributions (e.g., brightness, location) that roughly mimic the real astrophysical objects of interest for each particular survey (mostly supernovae, for the above examples). Typically, the MC method was then used to simulate light curves for the transient of interest, and then the derived detection efficiencies were applied.

The take-away message is that, for Rubin Observatory to best serve a broad section of the transient community, the simulated population of fake sources need only be representative of true astrophysical sources in a bulk sense, in terms of their brightness and location (i.e., plant more faint sources than bright, and more in high surface-brightness areas than isolated regions). The simulated fakes do not need to accurately represent astrophysical transient types, colors, redshifts, durations, light curves, etc., or be planted in sequential images in a correlated way to represent real light curves. That aspect of the analysis is best left to the user to handle during the MC simulation stage for their particular transient type.

A.2.1 Sloan Digital Sky Survey II (SDSS-II)

In order to calculate the occurrence rates of Type Ia supernovae (SNe Ia) from the SDSS-II, [20] generated a realistic sample of SNe Ia and injected fake point sources into the images as part of the live data processing pipeline to discover SNe Ia. Additional simulations to evaluate on how often the fakes were recovered by the end-to-end SN Ia discovery pipeline were then required to evaluate how assumptions about the simulated population (e.g., the distribution of light curve stretches) contributed to the final uncertainty in the derived rates [5]. The final form of their derived detection efficiency for SNe Ia was $\epsilon(z) = (0.78 \pm 0.01) + (-0.13 \pm 0.14)z$, within which is captured assumptions about the true relative fraction of each SN Ia subtype, such as the under/over-luminous 91bg/91T-likes [5]. The detection efficiency, ϵ, contributed to the final volumetric rate, $r_V = N/\widetilde{VT\epsilon}$, where N is the number of SNe Ia detected, and $\widetilde{VT\epsilon}$ is the product of the effective survey volume, time, and detection efficiency.

A.2.2 Dark Energy Survey (DES)

In order to determine the SN Ia detection efficiency as a function of redshift, the DES team used a method very similar to SDSS-II: fake sources were injected into their live data, which was run through their real-time `DiffImg` pipeline used to detect transients [17]. This process also started by simulating a realistic sample of SNe Ia, with parameters such as light curve stretch, host offset, and subtype drawn from established underlying distributions, and then used a Monte Carlo simulation of many more SN light curves, combined with the detection efficiencies for their fakes, to determine the SN Ia detection efficiency as a function of redshift.

A.2.3 A Canada-France-Hawaii Telescope (CFHT) Cluster Survey

In order to calculate the occurrence rate of SNe Ia in galaxy clusters for a CFHT imaging survey, [21] performed DIA to detect SNe in real time, but the fake injection was done separately. Simulated point sources were implanted into a representative subset of their images, and detection efficiencies calculated as a function of the relevant parameters for this survey: apparent magnitude, image quality, and focal plane location (because because the survey used single pointings with only small dithers). Their expression for the rate of Type Ia supernovae per unit stellar mass is $R_{\rm Ia} = (N_{\rm Ia}/C_{\rm spec})/(\sum_{j=1}^{j=N_{\rm img}} \Delta t_j M j)$, where $N_{\rm Ia}$ is the number of SNe Ia discovered in the survey, $C_{\rm spec}$ is the spectroscopic confirmation rate (determined separately), the denominator's sum is over all images of the survey, M_j is the total stellar mass within the image, and Δt_j is the control time for SNe Ia of that image. The control time is expressed as $\Delta t = \int_{t_1}^{t_2} \eta(m(t))dt$, where $m(t)$ is a SN Ia light curve, η is the detection efficiency, and the integration limits are the survey's temporal boundaries.

The Monte Carlo method was then used to calculate $R_{\rm Ia}$ for the survey many times while sampling over a realistic distribution of SN Ia light curve properties for $m(t)$ and the errors in $N_{\rm Ia}$, M_j, and eta. During this MC, an *effective* detection efficiency was used, which accounts for the possibility that the simulated SN Ia was detected in the previous two images: $\eta = \eta_j - \eta_j\eta_{(j-1)} - \eta_j\eta_{(j-2)} - \eta_j\eta_{(j-1)}\eta_{(j-2)}$ (as was appropriate for this surveys monthly cadence). The final result was quoted as the median of the MC rates with 1σ errors. A similar methodology was applied to this survey's cluster SNe II in [11].

A.2.4 Palomar Transient Factory (PTF)

The PTF covered 8000 square degrees with a three-to-five day cadence and generated over 1 PB of data. As detailed by [8], inserting fake sources into all of these images was both impractical and unnecessary. Instead, they chose a single representative *field* and planted fake sources in all images of that field. The fakes were given a uniform distribution in apparent magnitude, distributed in each image such that most of them are located within a galaxy, and then the PTF detection efficiencies were determined as a function of the apparent magnitude, the local surface brightness, and image parameters such as FWHM, airmass, moon illumination fraction, and sky background. These detection efficiencies were used to derive the volumetric rate of normal SNe Ia [10], of Ca-rich transients by [9], and of tidal disruption events (TDEs) by [13]. However, note that some rates analyses for TDEs have used aperture photometry and not difference imaging, and thus did not need fake injection for difference-image detection efficiencies (e.g., [12]).

A.3 Cosmology

The selection function for Type Ia supernovae that are used in dark energy analyses must be well constrained.

A.4 Active Galactic Nuclei

Active Galactic Nuclei (AGN) are powered by a supermassive black hole in the center of a galaxy, surrounded by a gas disk. Their energy output is non-thermal emission from X-ray through to mid-infrared, including emission lines in the optical spectrum, and many (or most) AGN exhibit optical variability. As such, they appear as a variable point source in the cores of galaxies. Many studies of AGN use aperture photometry on direct images, and not difference imaging, because it is desirable to have the *entire* flux of the AGN's point source, not just the flux in its variable component.

AGN samples have typically identified using spectroscopic emission lines, optical colors, or by looking for excesses of radio, X-ray, or mid-infrared emission, but selection by optical variability is also an option and in particular it may be better at including low-luminosity AGN, as described by [e.g., 24, 25]. The detection method of [24] uses image subtraction for the initial detection of AGN candidates, mainly because difference imaging was already done to find supernovae in the survey. Aperture photometry is then performed on the candidates, and a variability threshold applied to form their final sample for spectroscopic follow-up. [25] skip the difference imaging step and use aperture photometry and a statistical analysis true variability to identify AGN. Neither use the injection of fake point sources to evaluate their detection efficiencies, and instead use objects with spectra and/or X-ray detections to estimate their completeness. However, using spectra or X-ray to characterize the sample selection function of AGN identified by optical variability might not be possible in the LSST era, when optical variability becomes a more efficient and prolific way to discover a AGN [e.g., 2], generating significantly larger, lower-luminosity, and/or higher-redshift samples, for which spectroscopic confirmation is more difficult.

Furthermore, TDE and SNe also occur in the cores of galaxies [e.g., 19, 16], causing contamination of the AGN samples, and quantifying the rate of missed transients in the cores of galaxies remains an open problem.

Considering the needs of the AGN, TDE, and SN communities suggests that a population of fake injected transients to characterize the difference-image detection efficiency of point sources in galaxy cores would be scientifically beneficial.

A.5 Variable Stars

A star's luminosity might exhibit intrinsic variability (e.g., RR Lyrae, Cepheids) and/or extrinsic variability (e.g., eclipsing binaries, exoplanet transits, or microlensing events). In uncrowded fields, using direct images and the total flux is preferable to difference-imaging analysis for scientific studies that aim to identify and characterize variable stars. However, in crowded fields such as the Galactic plane, identifying variable stars in difference images can be much easier because the difference image is not as (or not at all) crowded, compared to the direct image.

For example, the census of variable stars in crowded fields by [7] describes how difference imaging is used, although it does not appear that injecting fake sources or deriving detection efficiencies was needed for their analysis.

It also seems that difference-image detection efficiencies are not needed for microlensing studies, for which it is a common methodology to fit a PSF to every pixel of a difference image, concoct a "light curve", and then statistically assess whether it is consistent with the

Draft

expected shape of a microlensing event [18][1].

Despite difference-image detection efficiencies not playing a role in past variable star studies, they might still be needed for LSST analyses. For example, injecting new fake sources in crowded stellar fields might be useful for detection efficiencies for stars which are too faint to have a counterpart in the template, but whose variable component makes them detectable by LSST for a short while. This would apply to e.g., M-dwarf flares (a common contaminant in searches for young SNe) and microlensing events.

Simulating variability of stars that *are* present in the template image is beyond the scope of this document.

A.6 Moving Objects

All of the moving objects identified by LSST will first be detected as difference-image sources, and detection efficiencies would be useful for population studies. In epochs of non-detection, being able to obtain the detection efficiency at a predicted location (i.e., as a function of local surface brightness and image qualities) would be a useful quantity for science goals related to moving object populations. The probability of, e.g., a faint asteroid's chance alignment over bright galaxies is small, and so fake injected point sources in empty locations may be more useful for moving object science — these would be needed to simulate very high-redshift transients, as well.

A consideration of whether the injection of trailed sources is scientifically useful is left for other work.

B Requirements Review

B.1 Summary

There are no requirements related to producing or serving detection efficiencies for transient `DIASources` in difference images.

[1] Although [18] does mention that detection efficiencies would be used in a paper in preparation.

Draft

However, the OSS specifies that spuriousness characterization be done *"by insertion and recovery of artificial sources"*, and the DMSR specifies that *software* for fake injection is a deliverable of the DMS (§ B.5).

This suggests that use of artificial source injection to determine a detection efficiency matrix for transient `DIASources` would not require a large expansion of scope for DM.

As a final note, the DMS computational system is sized to accommodate the (re)processing of all images with fake sources implanted **(from KT)**.

B.2 Science Requirements Document (SRD, LPM-17)

The SRD [LPM-17], V5.2.4 (Jan 30 2018), does not contain any requirements related to detection efficiencies.

The SRD does describe how the Prompt pipeline should provide *"the fast release of data on likely optical transients will include measurements of position, flux, size and shape, using appropriate weighting functions, for all the objects detected above* `transSNR` *signal-to-noise ratio in difference images"*, where `transSNR` = 5 (page 41).

The SRD thus specifies that all sources with $SNR \geq 5$ in a difference image are considered "detected", and it is the efficiency (or completeness) of detected transients (i.e., new sources which were not present in the template) that this document is concerned with.

B.3 LSST System Requirements Document (LSE-29)

The LSST [LSE-29], V7.1 (Mar 5 2020), does not contain any requirements related to detection efficiencies.

B.4 Observatory System Specifications (OSS, LSE-30)

The OSS [LSE-30], V19.1 (July 30 2021) does not contain any requirements related to the determination of detection efficiencies.

However, it does contain several specifications related to characterizing the spuriousness (also referred to as the “real/bogus" parameter), completeness, and purity of LSST sources detected in difference images. See Appendix C for a description of these terms.

The "insertion and recovery of artificial sources" that are mentioned as a means towards achieving requirements 0351, 0353, and 0354 could also be used to determine detection efficiencies.

- OSS-REQ-0351, *Difference Source Spurious Probability Metric*, specifies that *“the Observatory shall develop a metric to characterize the probability of each reported difference source being spurious"* (Section 3.1.5.2.1.7.5). The discussion further clarifies that *“the performance of this metric will be assessed by simulations, by insertion and recovery of artificial sources, and comparisons to ground truth where known"*. OSS-REQ-0351
- OSS-REQ-0352, *Difference Source Sample Completeness*, specifies that *“for each visit, the Observatory shall estimate the detected difference source sample completeness and purity as a function of the spuriousness metric threshold cut"* (Section 3.1.5.2.1.7.6). The discussion further clarifies that *“this information will aid the end users in selecting the spuriousness threshold appropriate for their particular science case"*. OSS-REQ-0352
- OSS-REQ-0353, *Difference Source Spuriousness Threshold - Transients*, specifies that *“there shall exist a spuriousness threshold* `T` *for which the completeness and purity of selected difference sources are higher than* `transCompletenessMin` *(90%) and* `transPurityMin` *(95%), respectively, at the SNR detection threshold* `transSampleSNR` *(6). This requirement is to be interpreted as an average over the entire survey"* (Section 3.1.5.2.1.7.7). As for OSS-REQ-0351, the discussion further clarifies that *“the performance of this metric will be assessed by simulations, by insertion and recovery of artificial sources, and comparisons to ground truth where known"*. OSS-REQ-0353 transCompleteness-Min transPurityMin transSampleSNR
- OSS-REQ-0354, is similar to OSS-REQ-0353 but defines a threshold for moving objects. OSS-REQ-0354

Regarding spuriousness, the discussion for OSS-REQ-0351 further describes that the *“spuriousness metric be prior free to the extent possible. For example, while it may make use of information from the source and image characterization (e.g., comparison of source to PSF morphology), as well as the information on the Telescope and Camera system (e.g., ghost maps, defect maps, etc.), it will not use any information about the astrophysical neighborhood of the source, whether it has been previously observed or not, etc. The intent is to avoid introducing a bias against unusual sources or sources discovered in unusual environments"*.

B.5 Data Management System Requirements (DMSR, LSE-61)

The DMSR [LSE-61], V9.0 (Feb 12 2021) does contain a few statements relevant to detection efficiencies.

- DMS-REQ-0097, *"Level 1 Data Quality Report Definition"*, specifies that the Data Management System (DMS) *"shall produce ... indicators of data quality that result from running the DMS pipelines, including ... detection efficiency for point sources vs. mag for each utilized filter"* (Section 1.3.14). However, DMS-REQ-0097 is derived from OSS-REQ-0131, *"Nightly Summary Products"*, and is clearly intended to produce a nightly summary of the general performance of the observatory and the DMS, not to provide the scientifically useful detection efficiencies that are the topic of this document. DMS-REQ-0097
- DMS-REQ-0009, *"Simulated Data"*, specifies that *"the DMS shall provide the ability to inject artificial or simulated data into data products to assess the functional and temporal performance of the production processing software"* (Section 3.2.1). This requirement is derived in part from OSS requirements 0351, 0353, and 0354 discussed above. DMS-REQ-0009

B.6 Data Products Definitions Document (DPDD, LSE-163)

The DPDD [LSE-163], V3.6 (Dec 17 2021) is not a requirements document but does have some information relevant to detection efficiencies.

Table 1 lists a single float is also reserved for SNR in the DPDD's `DIASource` table, defined as the *"signal-to-noise ratio at which this source was detected in the difference image"*.

Table 1 also lists a single float is reserved in the `DIASource` table for the spuriousness parameter. The DPDD's definition and description of spuriousness match the OSS requirement (i.e., this is a direct flow-down of OSS-REQ-0351). OSS-REQ-0351

The `DIASource` catalog table has three other relevant parameters listed that might be scientifically useful for analyses involving detection efficiencies:

- `psLnL [float]` – *Natural log likelihood of the observed data given the point source model.*

This represents the probability that a detected source is a point source; detection efficiencies would not apply to non-point sources.

- `fpBkgd [float]` $\mathrm{nJy/asec^2}$ – *Estimated background at the position (centroid) of the object in the template image.* This will be useful to provide detection efficiencies in difference images as a function of the background at that location.
- `fpBkgdErr [float]` $\mathrm{nJy/asec^2}$ – *Estimated uncertainty of* `fpBkgd`*.* Useful in the same way as `fpBkgd`.

B.7 Data Management Science Pipelines Design (LDM-151)

LDM-151, V4.3 (Nov 10 2020), details the scientific design and implementation of the requirements set by the SRD, OSS, and DMSR, and the generation of the data products described in the DPDD.

Section 3, *"Alert Production"*, states that *"In this document we do not address estimation of the selection function for alert generation through the injection of simulated sources. Such a process could be undertaken in batch mode as part of the DRP."*

However, LDM-151 does make two relevant statements about the `spuriousness` parameter which describe how the real-bogus algorithm will likely *"be based on a trained random forest classifier ... conditioned on the image quality and airmass"* (Section 3.2.4) and that it *"may use machine learning on other measurements or pixels"* (Section 6.7.2).

C Spuriousness (Real/Bogus)

The completeness, purity, and spuriousness of samples of detected sources are all related to detection efficiencies.

When source detection is performed on a difference image, both true astrophysical sources and artifact sources can be detected with $SNR > transSNR = 5$ and become `DIASources`. Examples of artifacts include, e.g., sources caused by telescope hardware, or incompletely removed cosmic rays (despite the fact that cosmic rays are real astrophysical messengers).

Spuriousness ($\mathcal{S}$): a parameter, typically between 0 and 1, which represents the probability that a source is astrophysical (values closer to 1) or artifact (values closer to 0).

The spuriousness is sometimes also called the *real/bogus score*. It is typically determined by machine learning on many sources that were pre-classified as astrophysical (real) or artifact (bogus), and usually considers factors like the shape of the point source in the difference image (e.g., symmetry, sharpness).

Spuriousness Threshold ($\tau_{\mathcal{S}}$): a lower limit on spuriousness, $\mathcal{S}$, which can be applied to `DIASources` to separate the "real" from the "bogus", and to thereby obtain a subset of `DIASources` with a lower fraction of artifacts.

The following definitions describe the classification of detected `DIASources`, which were either truly real (astrophysical source) or truly false (artifact), as real or bogus by an applied spuriousness threshold $\tau_{\mathcal{S}}$:

True Positive (TP)	A detected astrophysical source that was classified as real.
False Positive (FP)	A detected artifact that was classified as real.
True Negative (TN)	A detected artifact that was classified as bogus.
False Negative (FN)	A detected astrophysical source that was classified as bogus.

Completeness ($\mathcal{C}$): the fraction of all detected astrophysical sources which are classified as real. This is also called the true positive rate, $TPR = \frac{TP}{TP+FN}$.

Purity ($\mathcal{P}$): the fraction of all detected sources that were classified as real which are astrophysical in nature: $\frac{TP}{TP+FP}$. This is also called the *precision* or the *positive predictive value* (PPV) of a survey.

The false positive rate[2] is the fraction of all detected sources that were classified as real but which are artifacts: $FPR = 1 - PPV = \frac{FP}{TP+FP}$.

The relationship between the completeness and purity of a sample of `DIASources` classified as real can be traced out by varying $\tau_{\mathcal{S}}$.

In Figure 1 we show, as an example, the relationship between the false positive rate (FPR;

[2]Note that in some published literature the false positive rate is instead defined as the fraction of all detected artifacts that were classified as real, $FPR = \frac{FP}{TN+FP}$.

purity) and the missed detection rate (MDR; completeness) for different types of source classification (i.e., real/bogus) algorithms from the Palomar Transient Factory [PTF; 1]. This relationship is formally known as the Receiver Operating Characteristic (ROC) curve when it is plotted as the true positive rate *vs.* the false positive rate for a given `transSNR`.

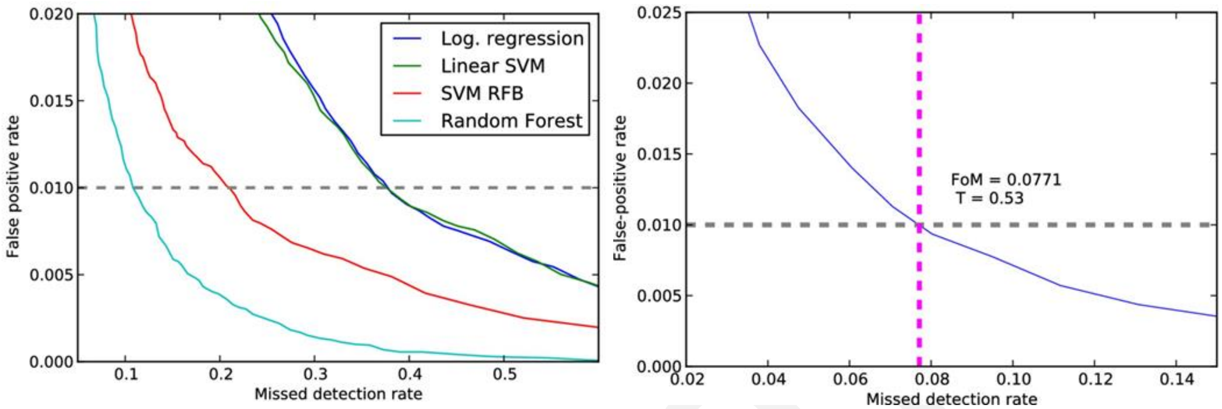


FIGURE 1: *Left:* An example of the relationship between the false positive rate (FPR; purity) *vs.* the missed detection rate (MDR; completeness) for different types of source classification (real/bogus) algorithms considered by the Palomar Transient Factory [1]. *Right:* The relationship between FPR and MDR for the RB2 (real-bogus version 2) classifier (blue line) developed by the PTF and introduced in [1]. Dashed lines show how $FPR = 0.01$ is achieved with a spuriousness (real-bogus score value) threshold of $\tau = 0.53$, which results in $MDR = 0.077$.

Characterizing spuriousness requires knowing which of the detected `DIASources` are astrophysical (truly real) and which are artifacts (truly bogus). Synthetic source injection is best for simulating astrophysical (real) sources because artifacts (bogus) can be a very diverse group; source injection cannot reveal which `DIASources` in an image are artifacts. Instead, characterizing spuriousness is achieved by building a training set of point sources that have been confirmed as astrophysical and artifact, using that training set for the spuriousness ($\mathcal{S}$; real/bogus) assignments by the machine learning algorithm. Then, since $transSNR \propto m$, the ROC curve yields $C(m, \tau_{\mathcal{S}})$ and $\mathcal{P}(m, \tau_{\mathcal{S}})$.

D Options Regarding Detection Efficiencies

The options for DM to participate in generating detection efficiencies, $\eta(\vec{P})$, are listed and discussed in terms of scope, risk, requirements, and science.

D.1 Do Nothing

In this scenario, the data from any fake injection that is done in order to meet the requirement to characterize the spuriousness is not made available, but the science community would have access to the *software* for fake injection.

Scope – No expansion of scope.
Risk – A moderate risk in that multiple user groups may then need to perform fake injection, leading to redundant reprocessing of the data and a computational strain on the resources.
Requirements – Does not violate or fulfill any requirements.
Science – This option would negatively impact science results based on transient phenomena, one of the four pillars of LSST science. The need for computationally intensive processing would force multiple teams to compete, and might limit the number of individuals or teams who could successfully derive detection efficiencies, and thus limit the scientific applications.

D.2 Make Available the Fakes Injected for Spuriousness Characterization

In this scenario, the data generated by the injection and recovery artificial sources in order to meet the requirements to characterize the spuriousness parameter is made available so that users may calculate detection efficiencies. For example, a `DIASource`-like catalog for the fake injected point sources, which users could bin by the $\vec{P}$ relevant to their science and generate $\eta(\vec{P})$. Since the current OSS requirements are to characterize the relationship between τ_S and sample completeness *only as a function of visit image qualities* for `DIASources` with SNR>5 (§ B), this scenario does not guarantee that these artificial sources will be adequate for all science use-cases.

Scope – Possible expansion of scope to make available the fake-source catalogs.
Risk – Minor risk, if the fake sources do not adequately cover $\vec{P}$, for the same reason as in § D.1.
Requirements – Does not violate or fulfill any requirements.
Science – Allowing users to build detection efficiency matrices from the same fake sources as are used to characterize spuriousness would enable at least some scientific analyses.

D.3 Ensure the Fakes Injected for Spuriousness Characterization Meet Science Goals

This scenario is similar to that in § D.2, except the fake sources that are injected and recovered in order to meet the requirements to characterize the spuriousness parameter are scientifically validated to cover the parameters needed for scientific analyses, $\vec{P}$, as listed in Table 1.

Scope – Minor expansion of scope to validate the artificial sources cover an adequate range of parameter space, $\vec{P}$, and to make available the fake-source catalogs.
Risk – No risk.
Requirements – Does not violate or fulfill any requirements.
Science – Allowing users to build detection efficiency matrices from a scientifically-validated set of artificial sources would enable more scientific analyses.

D.4 Generate and Provide Detection Efficiencies

This scenario takes that of § D.3 one step further, and has the DMS generate and provide the detection efficiency matrix, $\eta(\vec{P})$, as a scientifically validated and verified data product.

Scope – Moderate expansion of scope to generate $\eta(\vec{P})$.
Risk – No risk.
Requirements – Does not violate or fulfill any requirements.
Science – Enables scientific analyses for all that need detection efficiencies.

D.5 Inject Fakes during Prompt or Data Release Processing?

Here is considered three possible points in the data processing where the fake injection could be performed: during Prompt processing (§ D.5.1), on an intermediate timescale between Prompt and Data Release processing (§ D.5.2), and during DR processing (§ D.5.3).

D.5.1 During Prompt Processing

Inject fake sources into the live data which is processed within 60 seconds of image readout (OTT1). With this option, fake sources would be injected *on the fly into every new visit image from*

Draft

the telescope processed with DIA (or into the template image as negatives). This may seem like an extreme option to propose, but as discussed in § A.2 some previous transient surveys have injected fakes into their real-time pipelines, so we consider it here as well.

Scientific Motivation for Prompt Fakes – Surveys that plant artificial sources into live data processing typically use realistic light curves that represent their target population (e.g., color, duration, and location), and inject the fakes into sequential images in order to simulate real transients. The objective of this level of real-time injection is usually not just to characterize the detection efficiency, but also biases in the survey's classification algorithms and/or follow-up programs. Simulating fake sources that represent *real transient light-curves* in sequential images and in different filters in a realistic way is *not* being proposed here. Furthermore, most of the scientific analyses that require detection efficiencies, such as occurrence rates and population studies, would be done with the DIA products from a data release, and not the prompt products. A continually-updated detection efficiency matrix, $\eta(\vec{P})$, that incorporates data from fakes injected during Prompt processing does not have a strong scientific motivation.

Interference with Astrophysical Sources – In this scenario, fake sources would be planted into new images in a manner that samples the range of parameters for $\eta(\vec{P})$. This process would be comprised of three steps: (1) identify coordinates where the fakes should be planted, (2) fake injection into the image, and (3) bookkeeping for the fakes to ensure they do not contaminate the Alert Stream or the `DIASource` catalog. Fakes should not be injected at random locations because it is important to sample regions with higher surface brightness and to avoid the locations of known `DIAObjects`. If 1000 fakes are assigned random locations and injected into a 3.2 Gp image, and we assume that image has 10000 (randomly-distributed) true sources in it, then the probability that none of the fakes are coincident with one of the true sources is 0.9968. However, over a full night of 1000 visits, the probability that none of the fakes ever landed on a true source in any image is 0.0437, and it is most likely ($P = 0.2218$) that 3 fakes would have interfered with true sources.

Benefits and Drawbacks of Prompt Fakes from a DM Perspective – Two of the benefits (to LSST DM) of fake injection during Prompt processing are that (1) it would negate the need for separate re-processing of all fake-infused images, which could increase the total computational budget by up to a factor of 2, and (2) it could offer real-time feedback on evolution in the survey's completeness or purity, which might be useful — however, real-time feedback is not a necessary component of the DMS and could be obtained via the spuriousness parameters, as completeness and purity correlate mainly with bulk image properties. Two of the main drawbacks of planting fakes into "live" data are that (1) only a small number should be planted so as to minimize the risk of interference with real phenomena and (2) the additional steps of simulating, planting, and verifying fakes must be included in the computational budget for Prompt processing, which completes within 60 seconds for every new direct image and is already tightly constrained.

Scope – An expansion of scope in terms of FTE work hours and computational resources.
Risk – A risk to the DMS by adding three steps to the 60-second processing budget and potentially interfering with the completeness and purity of `DIASources`.
Requirements – Does not violate any requirements (and is not necessary to meet any requirements).
Science – This option would provide scientifically useful detection efficiencies, however, it may compromise science results if it interferes with the completeness and purity of `DIASources`. As there would be a limit on the number of fakes injected into every image, and restrictions that those fakes be away from most true transients and variables, this method would not provide the *best* characterization of the survey's detection efficiencies.

D.5.2 On an Intermediate Timescale

As a compromise between injecting fakes during Prompt Processing (above) and during Data Release Processing (below), fakes could be injected and recovered on a intermediate timescale (e.g., daily, weekly, monthly). There would be no need to reprocess *every* image because the goal is to build up a detection efficiency model as a function of parameters like host background, seeing and airmass. This pipeline could include only images in the parameter space where additional fakes are required. However, as with the proposed option to do fake injection during Prompt processing, there is no science case (or internal use-case) that requires the detection efficiencies updated in real time.

Scope – An expansion of scope in terms of both FTE and computational resources of the DMS.
Risk – A risk to the DMS (increasing the amount of processing done in between DRs).
Requirements – Does not violate any requirements.
Science – This option would enables rates analyses on the Prompt data products, but these analyses are more likely to be done on the DR data products anyway, so it is unlikely that this option opens the door for any new — or otherwise inaccessible — science.

D.5.3 During Data Release Processing

The benefits of implanting artificial sources into images during the DIA which occurs as a part of the annual DR processing is that fakes can be injected (1) only in locations where there were no detected `DIASources` and (2) in all images without increasing the total computational budget by any more than is required to inject the PSFs. As described in Section B, it is likely that fake injection would be done as part of DIA during DR processing anyway, in order to characterize the spuriousness parameter. This option is only be adding a step to ensure that the fakes are injected in a way that samples the parameter space $\vec{P}$ (Table 1), as needed to use the fakes for detection efficiencies. These DR-derived detection efficiencies could be used on the Prompt data products for the following year.

Scope – A mild expansion of scope in terms of FTE, but potentially not in terms of computational resources.
Risk – No risks.
Requirements – Does not violate any requirements.
Science – Enables science with both the DR and Prompt data products.

D.6 Summary of Options

It would be scientifically useful – with only a potential small increase in scope, if any – to ensure that the artificial sources implanted to characterize the `DIASouce` spuriousness parameter sample the parameter space needed for scientific analyses involving detection efficiencies, $\vec{P}$ (e.g., Table 1), and to make the data from the injection and recovery of fake sources available to users so that they can build detection efficiency matrices $\eta(\vec{P})$. It would be even more useful to provide $\eta(\vec{P})$ as a scientifically validated data product. For both scenarios, doing the fake injection during the DIA which occurs as a prt of the annual Data Release processing both achieves the science goals and minimizes scope increase and risk to the DMS.

E Options for Fake Point Source Injection Techniques

There are a variety of ways in which fake sources can be simulated and injected into the images. Some options are more suitable for transients (§ E.1), and some for variable stars (§ E.2). The following is a precursory presentation of the options, some of which have been used in the science studies discussed in § A. Typically, artificial sources – positive or negative – are added to the direct image *before* that image enters the difference imaging pipeline, so that the detection efficiency captures the end-to-end pipeline efficiency for detecting difference-image sources. (This is why fakes are not typically injected into the final difference image prior to source detection).

E.1 Simulating New Fake Objects

This applies to point sources that appear where there was no point source in the template image, such as transients like supernovae, variable stars that are undetectable in their quiescent state, and moving objects (assuming they're slow-moving enough to not appear as trailed sources, which is a different problem not included in this study).

Artificial sources that represent new fake objects would be planted in and around galaxies in a way that samples the environments of known transients (and serves the use-cases of variable stars and moving objects projected on background galaxies), would adequately sample areas of open space where most moving objects, some variable stars, and high-z transients with undetectable hosts will be found, and also in crowded fields where many variable stars will be discovered.

Model PSF — A 2D model for the PSF is added to the direct image in order to simulate a new point source. The shape of the PSF is derived from known trends with, e.g., the focal plane location or airmass (DCR), and the brighter/fatter effect.

Clone-Stamping — A nearby star is cut out and rescaled, and used as the simulated point source. One of the main drawbacks of using clone-stamping with LSST images is that incorporating the brighter/fatter effect into the simulation requires either that a star which is both nearby and of a similar brightness be used or that a model component added to the clone star to appropriately change the shape for the simulated brightness. Another drawback of clone-stamping is that very sparse/crowded fields might not have enough nearby/isolated point sources to use.

To decide between model PSFs and clone-stamping will require some testing in order to properly assess their performance and load on the computational resources. However, since knowing the PSF very accurately is something the LSST DMS will already be doing, it seems likely that the Model PSF option should be easier.

E.2 Simulating Variability in Real Objects

This applies to point sources that appear in both the template and direct image, such as variable stars and AGN. In this case, artificial variability is added to an existing point source in the direct image. Extra steps would need to be taken to ensure any real, low-level variability does not affect the results.

Model PSF Variable Component — A 2D model for the PSF with the desired flux of the variable component of the source is added to the object's 2D profile. This might only be useful for probing small flux changes, as it would be difficult be consistent with the brighter/fatter effect.

Scaling-in-Place — Cutout the star, multiply its 2D profile by a scalar in order to make it brighter or fainter, and add it back to the image. This could be modified to account for the brighter/fatter effect by, e.g., convolving with a kernel that both applies the effect and the desired variability, instead of multiplying by a scalar.

E.2.1 Planting in a Template Image

Could there be situations in which, in order to simulate variability, adding a new point source to the template only, or to both the template and the direct image, is needed? For example, in crowded fields, the detection efficiency for variable components of stars that are faint in the template image might be difficult to accurately measure because faint stars are hard to detect and isolate in crowded fields. This a necessary step in applying either of the two above methods for injecting artificial variability in real objects. Thus, it might be necessary to simulate new faint stars in the template *and* the direct image, with some flux difference between them, in order to derive detection efficiencies that are not dominated by the brighter, more well-distinguished stars in a crowded field. This would add complexity to the issue and might further expand the scope of the proposed option to provide scientifically validated injected fakes.

Draft

Option	Scope	Risks	Requirements	Science
Do Nothing	no expansion	potential risks	does not violate or fulfill any	negative impact
Make OSS-Required Fakes Data Available...	possible expansion	potential risks	does not violate or fulfill any	enables some science
...and Ensure OSS-Req Fakes Cover $\vec{P}$...	minor expansion	no risks	does not violate or fulfill any	enables more science
...and Calculate and Provide $\eta(\vec{P})$ in DR*.	moderate expansion	no risks	does not violate or fulfill any	enables science

FIGURE 2: A summary of the options with evaluated criteria, based on § D.

F Summary, Open Questions, and Suggested Future Work

This study is not yet finished and there remain some open questions to address, and further work is likely needed in order for an informed decision about the options proposed.

Open Questions for DM-SST:
(1) Have DM's plans evolved away from what's in the documents (§ B)?
(2) Does this work constitute a DMTN? It is not very technical – yet. One option might be to take this to the TVS community for input and write a joint TVS-DM document, which incorporates the further study items below.

Further Study:
- establish the extent of the parameter space, $\vec{P}$
- evaluate the accuracy needed for $\eta(\vec{P})$
- does the mode of fake injection (§ D) matter for science?
- will LSST's DIA be good enough to simply inject fakes into difference images?