***Study team responses***

*2017 March 24*

*The following document repeats the review team and PDCO comments on the LSST study reports and joint white paper. The study team contributors have inserted responses to the review comments, with the JPL comments appearing in blue and the UW comments in green.*

-------------------------------------------------

2017 February 28

**LSST NEO Study Review Team Final Report**

**Introduction**

The Large Synoptic Survey Telescope (LSST) is currently under construction on Cerro Pachón in Chile, and is scheduled to begin survey operations in the early 2020’s. LSST promises to be a transformational survey project during its nominal 10-year lifetime, and has great potential to detect and catalog near-Earth objects (NEOs) along with other solar system objects. Due to LSST’s fundamentally different approach toward NEO detection from the current state of the art, a credible assessment of its discovery potential has been difficult. To better understand how LSST may contribute to efforts to detect and track large NEOs down to 140 meters, NASA commissioned a joint study involving LSST project personnel based at the University of Washington (UW), and independent dynamicists at the Jet Propulsion Laboratory (JPL).

NASA convened the LSST NEO study review team (hereafter referred to as “the review team”) to assess the study team’s initial study plans and final reports. The review team’s initial assessment of the study plan was delivered on 2015 December 18 in a document titled “LSST NEO Study Review Team Recommendations”. The study team’s final reports were recently provided to the review team, consisting of three documents:

* “A Joint White Paper Summarizing Results from Independent Simulations of LSST’s NEO Discovery Performance,” by Chesley and Ivezić
* “Projected Near-Earth Object Discovery Performance of the Large Synoptic Survey Telescope,” version January 19, 2017, by Chesley and Vereš
* “The Large Synoptic Survey Telescope as a Near-Earth Object Discovery Machine,” (v2.1; Jan 17, 2017) by Jones et al.

The review team assessed these documents, and submits this report as a record of its findings.

**Executive Summary**

The submitted study reports provide a comprehensive discussion of the expected performance of LSST and its MOPS software suite. Both studies provide detailed accounting of the potential sources of false detections, and demonstrate currently available methods for reducing this noise to a level that will be sufficient for the intended computational resources available to LSST. Realistic simulations of the expected detection rates and software performance further support the conclusions of both studies.

The review team commends the JPL and UW study teams on producing thoughtful and comprehensive reports of LSST’s estimated NEO discovery performance. It is notable that both studies converged on similar predicted results, and are in approximate agreement with recent published work such as Grav et al. (2016).

While the overall conclusions of the reports seem to be complete and well-justified, the review team identified one area of concern related to the treatment of false detection rates. A small number of other minor points regarding overall clarity and consistency of the reports were also noted, and are identified below.

**Treatment of false detections**

The JPL and UW teams went to considerable effort to address the key issue of false detections in LSST imagery and whether they would overwhelm the ability of MOPS to successfully link detections into tracklets, tracks, and orbits. The review panel is impressed by the UW study team’s use of real images from a similar camera (DECam) along with prototype LSST image processing software to get a better handle on the false detection rate. The review panel is also impressed with the effort that JPL made in conducting a full-scale test of the ability of MOPS to link detections into tracks, tracklets, and orbits in the presence of realistic false detections. These efforts go a long way toward allaying the concern that the proposed survey cadence for LSST might be incompatible with NEO detection. We offer a few comments here that the authors should consider addressing to strengthen the discussion in the paper.

The UW paper provides evidence that a large percentage of false detections are associated with real astrophysical objects (transients, poorly subtracted stars, etc). This suggests that the density of false detections will increase with increasing density of astrophysical objects and with increasing limiting magnitude of the science images. The authors should add some discussion of source density in the DECam test images and whether this small sample is representative of the entire sky that LSST will observe. The authors should also discuss whether the number of false positives should be scaled up to reflect the higher source density at the deeper limiting magnitude expected from LSST compared to DECam.

*In addressing this topic it is important to distinguish static sources from other types of astrophysical transients and asteroids. While it is true that we find a large fraction of the overall false detections to be related to static sources (such as poorly-subtracted bright stars), we believe these detections to be unimportant for estimating the performance of MOPS since they are the easiest false detections to reject either algorithmically or, in a more brute force scenario, by simply masking bright static sources. The false positive rates we quote all (after the initial raw estimate of ~1000/sq deg) take into account this rejection of static sources, and hence would not increase if the density of static sources was higher. There may be an increased effective survey area loss, depending on how well these rejection algorithms perform, but this is still likely to remain below 1% when averaged over the LSST survey area.*

*In the case of astrophysical transients, their number density will indeed increase with observation depth. With the extra magnitude of depth in LSST and the observed SNR^-2.5 distribution we see of DIASources, this would be an extra factor of 4 in detections. However, the extent to which this is a feature or a problem depends on understanding what fraction of our DIASources are true asteroids. Increased numbers of asteroids at LSST depths is clearly beneficial and is already accounted for in the computing power requirements, and should not be treated like false positives. Because our data likely contains some mixture of both true sources and false sources, our selected scaling factor of 1.97 (between DECam and LSST) approximates the scaling that would be seen if 50% of our detections were astrophysical but non-asteroids (e.g. supernovae), and 50% were asteroids.*

*Regarding the possibility that false detections could be elevated in high-density regions, JPL has found that removing all Galactic Plane proposal fields leads to only a 0.2% degradation in completeness. This point has been added to the JPL text on false positives.*

One aspect of the false detection analysis that is somewhat weak in both the JPL and UW papers is the discussion of spatially-correlated false detections. The JPL paper (section 2.5.2) refers to a paper by Slater et al. (2016) that claims spatially correlated false detections occurred in DECcam data at ~2/deg2 and were therefore negligible. However, the Slater et al. (2016) paper does not specifically address spatial correlation of the false detections. It estimates the number of false positives associated with very bright stars, which is one source of spatially-correlated false detections. It does not estimate how often false positives from bright stars might line up to form false tracklets. The quoted number (2/deg2) appears to come from the accompanying UW paper, not the Slater et al. (2016) paper.

*We have added a new analysis of tracklets near bright stars to address this issue. Section 3.3.3 - “Spatially Correlated Transients” of the UW paper now contains a figure showing “stacked” tracklets surrounding bright stars, which would exhibit correlated positions and alignments if bright stars were causing false tracklets via diffraction spikes, etc.*

*The JPL Report has changed the citation to point to Jones et al. 2017.*

The UW paper addresses the issue of spatially-correlated false detections in one paragraph in Sec 3.3.3:

*We also investigated whether DIASources from different difference images are correlated in focal plane coordinates. A large number of such correlated detections could result in a substantial increase of tracklets and tracks. We analyzed 4 visits with 60 CCDs each, and found 24 DIASources that match to within two pixels. We did not find any correlation at larger radii. These 24 DIASources show tendencies for certain CCDs but the counts are too low to make a robust conclusion. Visual inspection shows that many are near parts of an image where a defect (such as a cosmic ray, bad column, bleed trail, etc) had been interpolated over, though for some the cause is unclear. The implied density of correlated DIASources is about 2.3 deg2, rendering this effect relatively unimportant.*

This paragraph is difficult to interpret and provides insufficient justification to ignore spatial correlations among false detections. We suggest that the UW team expand on this paragraph to better describe what they did. The authors should describe the contents of this test data more fully including whether these 4 visits of 60 CCDs each were taken from the Slater et al. (2016) data from DECam or from some other source. The authors should provide some justification that the small sample of images is representative of the full sky that will be observed by LSST.

*We have added some additional description of this calculation. We also hope that our additional treatment of false tracklets (discussed above) will provide greater confidence that correlated detections do not contribute significantly to tracklet rates, since it is a more robust test.*

*The primary goal of this work is to test the components of the LSST system that are common to all fields, rather than testing every region of observing parameter space. As an example, if our software was not able to properly account for correlated noise, that effect would be present in all observations, regardless of position in the sky. Similar arguments apply to the treatment of bright stars and sensor defects, and the quality of the sensors themselves. While it is possible that some feature of particular sky regions will be problematic, our focus is on showing a very low level of common-mode failures that would affect the survey as a whole.*

Although both reports provide some discussion of the applicability of machine learning to help reduce the false detection rate, it is not certain that the actual LSST data will be analogous to previous studies. For example, the UW report (p. 10) uses the results from Goldstein et al. (2015) as a metric for the ability of machine learning to improve false detection rejection. However, as shown in Figure 4 of that paper, while the diagnostic metrics highlighted there as most important do work well at SNR~10, at SNR~5 there is significant overlap between the real and false populations. As such, the potential gains from machine learning for lower SNRs will not necessarily be comparable to the gains found for higher SNRs. This may affect the extrapolations of full system performance, and the total numbers of false detections.

*ML was not assumed for any aspect of study. A clarification of this point was added to the UW paper.*

The UW report (p. 53, and elsewhere) states that 400 false detections per square degree is a conservative upper limit. However, this is based on the DECam images, which (although likely comparable to expected image quality on clean sky) may not be a good analog for high-density or otherwise noisy regions that occupy a small region of the total sky but can require a significant fraction of the computing resources. For general estimates of system performance, 400 false detections / deg2 is likely typical, but should not be considered 'conservative'.

*For reasons listed in sections 3.3.1 and 3.3.2 in the UW paper, as well as in p. 53, the use of the word “conservative” appears justified (the true false positive rate is probably \*much lower\* than the adopted value because in all ambiguous estimates the worst case scenario was adopted).*

**Minor points of consistency or clarification**

* The UW paper references deriving the noise model from a subset of 540 DECam images (section 3.3). The authors should clarify whether this larger set of 540 visits refers to the Slater et al. (2016) paper, and also whether the 15+5 visits used here are a subset of the 540 images, or are derived from a different data set.  
    
  *The Slater et al. (2016) work used the same 15+5 datasets as is reported here. We have added a clarification that the 540 visits are a superset of the ones analyzed in this work.*
* Although stationary transient noise is modeled, there is no discussion of the impact of moving transient artifacts. In particular, scattered light and diffraction spikes from the planets and the largest Main Belt asteroids will move at solar system rates, but are difficult to predict and eliminate as effectively as artifacts from stars. Discussion of these noise sources, mitigation plans, and impact on the simulated survey (in particular the linking step) would be beneficial to the final report.   
    
  *The number and implied sky density of sufficiently bright asteroids to produce scattered light and diffraction spike artifacts is at least two orders of magnitude smaller than the sky density of asteroids down to LSST depth on the Ecliptic. A comment to this effect was added to the UW paper.*
* The JPL study implemented a floor of 10 mas relative astrometry (p. 30) but the UW study states that the maximum absolute astrometry will be 100 mas, and possibly 10x better using Gaia (p. 6). These statements appear to be contradictory.   
    
  *Indeed, this does appear as an inconsistency, but not one that would compromise the results of either study. The JPL assumed that Gaia substantially eliminates absolute errors on such large fields, which reduced uncertainties being fed into the linking engine. The UW report did not pass synthetic observations to a linking engine.   
    
  Work in progress by the LSST team using Gaia’s recent Data Release 1 shows that indeed absolute astrometric errors will be negligible compared to random errors for all but the brightest objects. A comment to this effect was added to the UW paper.*
* In the JPL report’s Figure 9, top left plot, the axis labeled "inclination" should say "eccentricity". In Figure 23, the blue histogram is not described.   
    
  *- Panel from Fig. 9 corrected.   
  - Legend and caption of Fig. 23 revised to clarify.*
* In Figure 24 of the JPL report, the method of plotting the data obscures the density of points in the plots. It is not clear that the period plot will properly model reality, as it does not reproduce the 2-hour spin period limit for real objects.   
    
  *The plots have been redone as density plots, which better shows the relative distribution of objects. We have cited Masiero et al. (2009), who posited a strength curve for smaller main-belt objects that allows faster rotation.*
* There appears to be an error in the implementation of equation 13 in the JPL report. Plugging in the numbers assumed by the authors gives 16,000 false positives per image, rather than the 650 quoted. This seems to be due to the fact that the authors divided by the factor of 𝜂 (SNR) in equation 13 rather than multiplying, resulting in an underestimation of the false positive rate by a factor of 25. It is possible that there is a typo in equation 13, however this is derived from source material (i.e. Kaiser 2004) that is not publicly available and thus cannot be checked. If possible, the report should use publicly available references or reproduce the methods used to derive equation 13. However, if this is not a typo, then this significant underestimation of false positive rate will likely have a significant impact on the rest of the simulations, and should be corrected.   
    
  *There was a lapse in the description of the variables. In this equation, the seeing must be in pixels rather than arcsec. With LSST's 0.2 arcsec pixels the effect of using arcsec instead of pixels is 0.2-2 = 25. The passage has been revised to eliminate the miscommunication.*
* There is an error in the construction of equation B24 of the UW report, and the resulting final equation giving the expected number of false tracks. At the top of p. 63, the report states that the velocity consistency criterion goes as *pv2*, instead of to the third power, as the velocity of the initial tracklet used to link to the other two is an assumed input. However, when this is propagated to equation B24, a factor of *pv3* is used, becoming the *vmax* -6 term. This results in equation B25 having a final relation of *vmax*-1.3 which would imply that the number of false tracks can be reduced to zero by simply setting the maximum allowed search velocity to an infinite rate of motion. If instead the *pv2* relation is used, then equation B24 will have a *vmax* -4 term, and equation B25 will have a final dependence of *vmax*0.7. This gives a more realistic result that as the maximum velocity searched increases, so too does the number of false tracklets.   
    
  *The perceived errors do not exist but we agree that explanation was not entirely clear. We have modified the UW paper (text after eqs. B24 and B25). In short, there are 1+2 velocity filtering steps so the final scaling in eq. B24 is indeed with v\_max^6. On the other hand, the numerical approximation given by eq. B25 is only valid around fiducial values and thus it cannot be extrapolated to the v\_max -> infinity limit.*
* It should be noted that filtering tracklets based on brightness can result in high interest objects being incorrectly removed. A magnitude variation of 1.5 magnitudes is not unheard of for NEOs. Increasing this to ~3 magnitudes would be safer, but removing the cut entirely would be best. A cut of the proposed type would remove objects with high light curve amplitudes, and also objects rapidly changing in brightness due to a close pass with the Earth, which are of significant interest.   
    
  *Presumably this comment is in reference to the discussion at p. 43 related to the higher SNR of artifacts. The reviewers point is well considered, but selecting the parameter for photometric consistency is beyond the scope of this work. The passage has been revised to acknowledge both the benefits and risks of enforcing photometric consistency, but now communicates a preference for only weak enforcement of photometric consistency, if at all.*
* The UW report asserts on p. 6 that astrometric accuracy will be improved by the application of the Gaia catalog. While this is true for high SNR sources, the total astrometric uncertainty on low SNR detections will be dominated by random noise, not catalog errors, and thus absolute astrometry for most of the sources of interest (newly discovered NEOs) will not be improved by an order of magnitude with better catalogs.   
    
  *A clarification that this statement pertains to the accuracy of absolute astrometric calibration, rather than the accuracy of absolute astrometry for individual sources, was added to the UW paper.*

**LSST NEO STUDY REPORT**

**Additional Comments from PDCO**

2017 March 7

**I. JPL Report**

Larger Points:

p. 15: Explain how realistic weather and seeing conditions are included in the OpSim.

*These are based on historical records from the observing location. The description has been expanded to address the question.*

p. 77: This needs a deeper explanation of how, and when, triples/quads can result in an orbit and a discovery. For instance, if a quad on one night provides 2 or 3 tracklets, what is needed on subsequent nights to generate an orbit? Since a quad *visit* doesn’t have to result in a quad *detection* in order to produce a tracklet (by the definition on p. 12), will quads be more robust to fading and light curve losses? These effects ought to be quantified at some level, maybe in a best-case scenario, to show whether they can offset the loss in completeness due to reduced sky coverage. The question of *when* discovery occurs should also be dealt with, because the possibility that an impacting object could be detected, but not discovered, before impact will be of concern.

*The question seems to arise from a misapprehension of the purpose and value of quads vs. pairs, as well as some misunderstanding of the linking process itself. The discussion has been substantially rewritten to hopefully eliminate the confusion and to emphasize that quads only lead to more reliable tracklets and not better orbits. A brief discussion has been added regarding the possibility that quads could produce reliable two-night tracks, and we now quantify the performance effect from requiring 3 out of 4 or 4 out of 4 detections of the two-nighters, and the dependence on the arc length. It is important to emphasize that these two-nighters are not likely to meet cataloging objectives in general.*

Smaller Points:

p. 5: Avoid saying that including other survey systems “increases LSST’s end-of-survey completeness.” It increases the completeness achieved by the total NEO detection system.

*Revised to clarify.*

p.8: Should say “has been estimated by the LSST Project to be capable of discovering” instead of “will discover”.

*Revised to clarify.*

p. 11: “...should already *be* discovered...”

*Corrected.*

p. 13: “...of detections *from two* or more...”

*Corrected.*

p. 30: Does the detection model take into account the higher local noise level in the spots where stationary sources have been subtracted in difference imaging?

*The JPL report now notes that m5 is assumed to be constant across the field. It also states “While sensitivity in a given image may be reduced in areas with higher background noise, the affected region is small, especially after bright source masking as described in Sec.~\ref{sec:artifacts}, and the performance impact is presumed negligible.”*

*When doing actual image processing, LSST carries the variance plane to track the noise level, so we do not end up with excess detections around stationary sources (i.e. the increased noise is tracked, but this does result in a loss of sensitivity around bright stars). Our estimate of 1% of area being lost to masking bright sources is a conservative estimate of how much area would be masked and should cover this area of lower sensitivity as well.*

*In our modeling, we assume a single m5 value across the focal plane (i.e. no areas of lower sensitivity). However, the JPL modeling discarded an extra 1% of the area, mimicking this loss of sensitivity around bright stars, and did not find a significant effect on completeness.*

p. 33: Figure 18—is this all asteroids or all detected/detectable asteroids? In which model?

*The plot shows rates for detected asteroids in enigma\_1189. Revised caption to clarify.*

p. 43: Symbol typos on this page. Also: evidently you didn’t consider the contribution of astrophysical transients to the false positive rate, whereas Jones et al. did. Should this be discussed?

* *Typographical errors corrected.*
* *JPL report clarifies (Sec. 2.5) that stationary transients are recognized as such by the LSST pipeline and these will not enter the MOPS pipeline. Jones et al. include astrophysical transients only in the sense that such phenomena could be present in the 400 detections per square degree that was derived from the DECam results.*

p. 49: It should be stated here for clarity that all results until section 3.7 are based on the assumption that LSST is working in isolation and that no NEOs are known at the start. Near the bottom of the page: not “affected by a factor of 1 percentage point”, merely “reduced by 1 percentage point”.

* *Clarification has been added as suggested.*
* *Text near bottom of page revised as suggested.*

pp. 52-53: Just “atop”, not “atop of”.

*Passage revised.*

p. 61: Figure 42 legend, FD = false detection? NS = ???

*The caption has been expanded to clarify the interpretation of the plot legend.*

p. 62, bottom: “Tho” isn’t a word.

*Corrected typographical error.*

p. 71: Can the results of the extended linking simulation be added/updated?

*This analysis could not be completed within available resources and schedule. We have added a passage stating that, "analysis of those products could not be completed within the schedule constraints of this study."*

p. 76: More accurate to say “...within a span of a couple of hours” rather than “an hour or so”.

*This is in reference to existing surveys such as CSS and Pan-STARRS, which rarely (if ever) produce tracklets longer than one hour. The passage has been revised to emphasize that this relates to past and ongoing surveys and not to our LSST simulations.*

p. 77: Should “the large ratio of tracks” be “the large fraction of false tracks?”

*Revised to clarify.*

p. 80: “Claimed” could be interpreted as a charged verb; suggest “found” instead.

*Agreed. Revised to say "they reported..."*

p. 82: Sounds like you are recommending eliminating the rate cutoff and including the deep- drilling and singleton fields in order to increase completeness; is that the intent? Clarify.

*It is not obvious what to clarify here. There is no suggestion of a recommendation in the passage, but rather just a listing of the several simulation differences and a mention as to which way the modeling difference would drive the final UW result.*

p.89: What’s the justification for increasing *F*disc to 0.8 and 0.9 at years 0 and 5? Is this consistent with what the UW team assumed?

*Our model of the other surveys was derived independently of (and in complete isolation from) UW. Nevertheless, the two models show excellent agreement. We increased the effective performance of the systems in the future to account for expected continuing improvements, specifically the planned upgrades to CSS and the anticipated initial operational capability of Pan-STARRS2. We have added a few words to explain this point.*

p. 90: The last sentence of section 3 is true, but the more relevant conclusion is that adding LSST to the existing NEO discovery system increases the completeness achieved in 2032 by 16 percentage points (i.e., from 61% to 77%) over what the existing system, with expected increases in efficiency, would achieve on its own. This should be stated explicitly.

*Revised as suggested.*

p. 91: It is somewhat misleading to say that adopting triples or quads “would dramatically decrease the NEO discovery rate”, since these approaches would be adopted only to remediate a serious failure of pairs. The correct statement would be that, if the pair approach for some reason failed and LSST were forced to go to triples/quads, it would not achieve the discovery rate initially (and erroneously) predicted for pairs.   
  
*It is not quite correct to state that triples or quads are considered only a fall back cadence in the event that pairs are ultimately unlinkable. There has been pressure from some in the NEO community for LSST to drop the two-visit-per-night cadence even before entering commissioning, due to pessimism that the linking problem can be managed. Even so, the sentence has been recast to state that, within the study hypotheses, the triples or quads have a lower performance than that of pairs, rather than stating that they would decrease the performance, which implies some sort of switch during the survey.*

p. 92: First bullet, 60% should be 61% to agree with earlier text. Second bullet, final sentence, sounds like a serious deficiency in the study—unless what you really mean is: “No NEO- enhanced cadence that we have tested significantly outperforms the baseline.”

* *Corrected to read 61%.*
* *Revised to clarify.*

**II. UW Report**

Larger Points:

pp. 8-9ff: The statement that false positives are “*assumed* not [to] be an issue” [emphasis added] is surprising. There is no actual test of linking performance in this report, except for the discussion of computational expense. The effects of false positives on the completeness and accuracy of the NEO catalog are not assessed. The analytic treatment of the false track rate in Appendix B stops just short of its goal, which is simply stated in equation (B25) as having been “determined numerically”; yet no numerical results are presented to support it. This is disappointing, as an opportunity for transparency is being missed. (The JPL team did assess linking performance, and they conclude that complete and accurate linking can actually be achieved. This supports the consistency of the final results from the studies. But they also demonstrated that assessing and ensuring the accuracy of the NEO catalog is far from trivial.)

*We agree that the choice of word “assumed” in that context was misleading. We have modified the UW report. We have also added a citation to the JPL report about linking performance.*

Smaller Points:

p. 2: There are significantly more than 6,000 NEOs with *D* > 140 m.

*Updated number of NEOs, used Harris reference but pieced out total # from figures/charts.*

p. 6: The sentence, “Discovering and linking...places strong constraints on the cadence of observations” implies that finding moving Solar System objects was a main driver of the baseline cadence, which is questionable. It should also be explained how realistic/optimistic/pessimistic the weather and seeing models in OpSim are.

*This statement is correct. The two visits per night strategy is primarily driven by moving objects (and supported by the desire to have two detections per night to handle false positives when searching for explosive astrophysical transients). Other science programs would instead prefer a single visit per night. Section 5.1.2 in the UW paper already states that the weather and seeing distributions used in OpSim are not models, but the actual multi-year measurements obtained at the Cerro Pachon site.*

p. 18: Given the steep scaling of tracklet number with false-positive density (Fig. 2), it seems like wishful thinking not to run any tests at values of the density higher than the expected one.

*That project was completed before the (upper limit for) false positive density was available. These tests will certainly be repeated several times before LSST operations start, as part of the commissioning of the LSST computing system and MOPS.*

p. 22: Does the detection model take into account the higher local noise level in the spots where stationary sources have been subtracted in difference imaging? The “model that emulates MOPS performance” needs to be thoroughly described. (Given the assumption on pp. 8-9, is the “model” simply that all real tracklets are found, all real tracks are linked, and all false tracks are removed?)

*The “detection linking” section of the completeness modeling in the UW report was modified to clarify the model that emulates MOPS performance, and to indicate that local m5 variations are not included in this model. (will also include effects of removing a CCD/Raft once those are available).*

p. 29: Is the “astro\_lsst\_01\_1016” OpSim survey the same as the “astro\_1016” survey used by the JPL group?

*Yes. The JPL report has clarified that it is using an abbreviated form of the name for this simulation.*

p. 36: The paragraph arguing that the biggest gain to completeness comes from extending the survey from 10 to 12 years reads, “For the case of PHAs and 30-day wide MOPS window, the completeness can be boosted from 65.6% after 10 years with minion\_1016 survey to 73.9% after 12 years with ‘extra ecliptic visits’ survey...” This is misleading on two counts. First, the figure for minion\_1016 should be 68.4% (the 30-day window figure from Table 2). Second, this is an apples-and-oranges comparison—extra ecliptic for 12 years *instead of* minion for 10, not minion for 10 years followed by extra ecliptic for 2 more.

*Added info on 12 year ‘baseline’ survey equivalent (astro\_lsst\_01\_1019) - see increase of 3.7% (vs 10 years of the same survey). Kept info on 12 year extra ecliptic visits survey as well (increase of 4%). Simplified concluding paragraph.*

p. 38: The title of section 5.4 isn’t quite right—you’re assessing the impact of non-LSST discoveries, both before and during LSST’s operating lifetime. Having done so, the completeness calculated is that reached by the entire NEO discovery system.

*The UW paper was revised to implement this correction.*

p. 39: The *H* < 22 criterion may be commonly used, but the *D* > 140 m criterion matches the Congressional mandate; there’s no rerason to criticize GMS for “redefining” the completeness limit.

*The word “redefined” was removed and text modified accordingly.*

pp. 39 & 42: On the former page: “...GMS results are fully consistent...with the results previously published by the LSST team...” On the latter: “All these recent results are...consistent...with even the earliest forecasts of LSST performance...” While the conference paper cited (Ivezic et al. 2007) does quote a 75% completeness figure for PHAs with *H* < 22, the more recent and highly cited arXiv:0805.2366 (version 4, 2014) asserts that the baseline survey—on its own, without existing surveys—will reach 82% completeness in 10 years, or, with an NEO-optimized cadence, 90% in 12 years. These figures have been quoted in public by the team many times.

*We agree that these somewhat higher values should be commented on. The following sentence: “The PHA completeness forecasts published in \cite{LSSToverview} are somewhat higher than reported here, primarily due to using different PHA populations.” was added to the UW paper.*

**III. Joint White Paper**

p. 1: Give total completeness (including other surveys) for NEOs here in the abstract, in addition to (or instead of) PHAs.

*Added NEO completeness metrics to abstract as requested.*

p. 3: The definition of when an object is catalogued precedes, and seems to be independent of, any discussion of the linking process. Can this be re-ordered so as to have a better logical progression? Re: first bullet (and on p. 5), avoid implying that the function of the other NEO discovery assets is to enhance the performance of LSST’s survey. LSST will (we expect) join the NEO discovery system, and it is the performance of the entire system that matters.

* *Added a clarifying sentence in the cataloging discussion pointing forward to the linkage discussion.*
* *On p. 3 and 5, recast description to avoid misinterpretation regarding the importance of contributions from other surveys.*

p. 5: The 2nd paragraph of section 3 contains the main results and requires some care. It would be helpful to cite the sections or page numbers in the main reports where these numbers appear. What is the “worst-case linking efficiency” that results in 55% completeness? It would be good here to reproduce Fig. 54 of the JPL report and Fig. 10 (left) of the UW report side-by-side. These figures indicate the total system completeness for NEOs should be approximately 75% after 10 years of LSST operation—this is clearer and more accurate than saying that it will “approach 80%”. It should also be pointed out that, even despite the differences in modeling LSST and the extant surveys, the figures show a robust result that the NEO discovery system should achieve an NEO completeness about 16 percentage points higher with LSST than with the current assets (with projected enhancements) operating alone.

*Rewrote and expanded paragraph in question. Added figure as suggested.*

p. 6: In the second bullet, give the results also for NEOs, and make it clear that this result depends on the assumption that existing surveys will continue to operate with improved sensitivities.

*Revised as requested.*