# The NEOCP in the Era of LSST

T. Wagg ,[1] M. Juric ,[1] and ... more

[1]*Department of Astronomy, University of Washington, Seattle, WA, 98195*

## ABSTRACT

[TODO: ] Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

## 1. INTRODUCTION

[TODO: ] Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim.

Corresponding author: Tom Wagg
tomjwagg@gmail.com

Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetuer at, consectetuer sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

Morbi luctus, wisi viverra faucibus pretium, nibh est placerat odio, nec commodo wisi enim eget quam. Quisque libero justo, consectetuer a, feugiat vitae, porttitor eu, libero. Suspendisse sed mauris vitae elit sollicitudin malesuada. Maecenas ultricies eros sit amet ante. Ut venenatis velit. Maecenas sed mi eget dui varius euismod. Phasellus aliquet volutpat odio. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Pellentesque sit amet pede ac sem eleifend consectetuer. Nullam elementum, urna vel imperdiet sodales, elit ipsum pharetra ligula, ac pretium ante justo a nulla. Curabitur tristique arcu eu metus. Vestibulum lectus. Proin mauris. Proin eu nunc eu urna hendrerit faucibus. Aliquam auctor, pede consequat laoreet varius, eros tellus scelerisque quam, pellentesque hendrerit ipsum dolor sed augue. Nulla nec lacus.

## 2. METHOD

In order to make predictions for the NEOCP in the era of LSST, we make simulated observations of a catalogue of solar system objects that takes into account currently known objects. We then use the `digest2` code to calculate NEO scores for each object and use these values to make predictions for the NEOCP. In the subsections below we explain each of these steps in more details.

### 2.1. *Hybrid Catalogue Pipeline*

Most studies that make predictions for LSST use a synthetic catalogue of solar system objects that doesn't account for prior observations (needs citation). In reality, we have already detected more than a million objects in the solar system and this number will continue to grow until LSST comes online. This means that, current predictions of detection rates will be inflated since a fraction of "new" detections may already be known. Therefore, for this paper we created "hybrid" catalogue that combines a synthetic catalogue with all known observations, whilst keeping the population distributions relatively unchanged.

We created the hybrid catalogue to be dynamic, such that we can run a single pipeline to merge in an updated
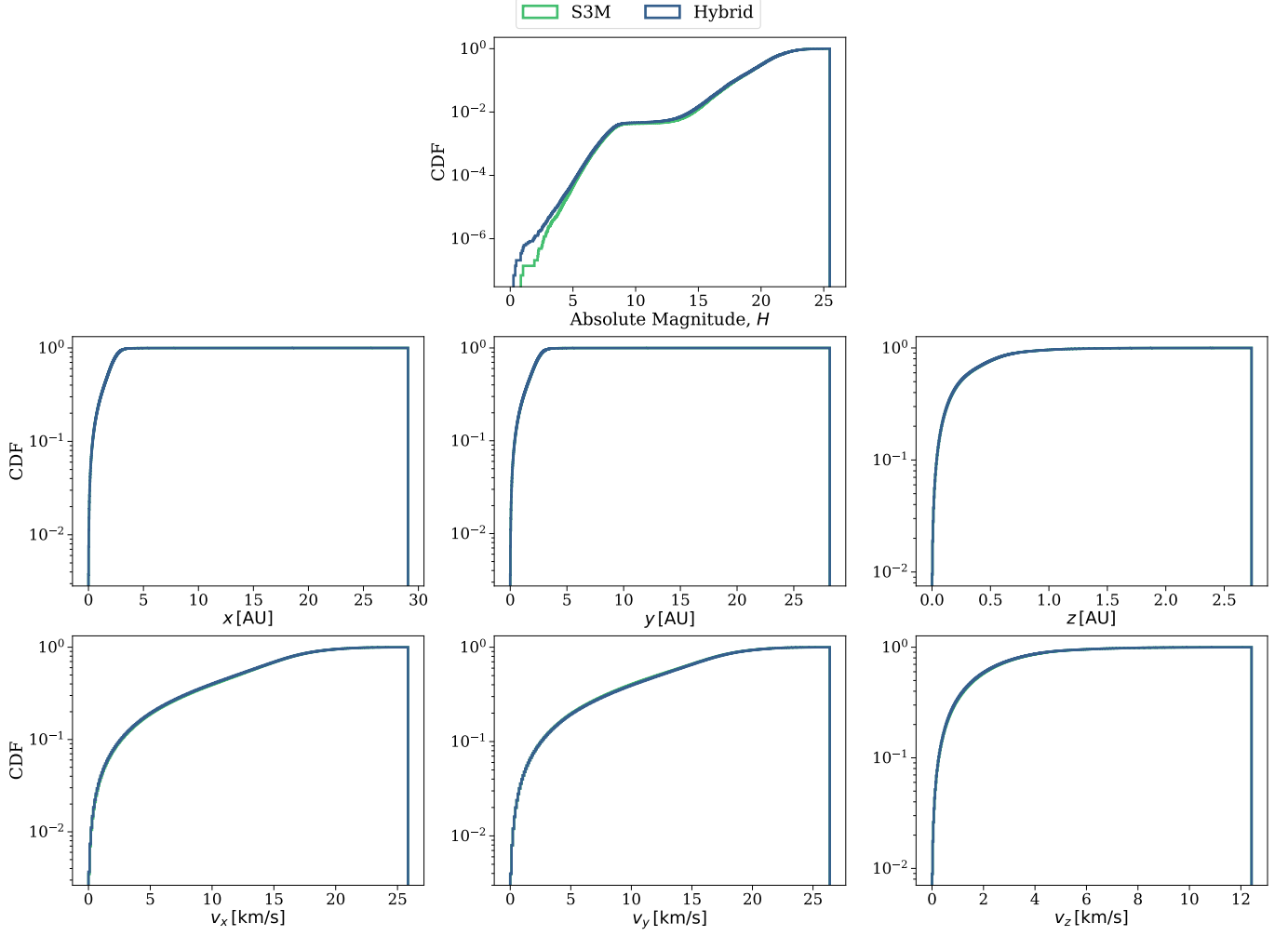
**Figure 1.** A comparison of the parameter distributions of S3M (Grav et al. 2011) and the hybrid catalogue we created.

version of MPCORB as more objects are discovered in the time until LSST comes online. All code to reproduce this hybrid catalogue is open-source and available on GitHub[1].

### 2.1.1. *Data preprocessing*

For the synthetic catalogue of the solar system with we use S3M, the Pan-STARRS Synthetic Solar System Model (Grav et al. 2011). We merge this synthetic catalogue with the latest version of MPCORB[2], a database of all currently known objects.

We use `OpenOrb` (Granvik et al. 2009) to convert both catalogues to Cartesian coordinates and propagate all orbits until the same date.

### 2.1.2. *Merging algorithm*

The general idea for the merging algorithm is to inject each object from MPCORB into S3M, replacing objects that are similar to those injected. An object's similarity is determined based on its position, $\vec{x}$, velocity, $\vec{v}$, and absolute magnitude (size), $H$.

We split each catalogue into bins of absolute magnitude linearly spaced from $-2$ to $28$ and perform the merge algorithm on each bin separately. For each bin we build a K-D trees for both catalogues based on the positions $(x, y, z)$ of objects. For every MPCORB object we query the S3M tree for the nearest 100 objects up to a maximum distance of 0.1 AU, excluding any that have already been matched to a different real object. From these remaining nearest neighbours, we select the S3M object with the closest velocity as the matched object. If there were no remaining neighbours, either because no synthetic objects were nearby or because all nearby objects had already been matched, then we directly add this real object without replacing a synthetic one.

To complete the merging process, we compile the matched object IDs and delete them from S3M. We then add the entirety of MPCORB to the remaining catalogue, resulting in a hybrid catalogue.

### 2.1.3. *Assessing quality of hybrid catalogue*

It is essential that the underlying distributions of the hybrid catalogue do not differ significantly from S3M so that we still accurately reproduce the solar system. In Figure 1, we show the distributions of the absolute magnitude and six orbital elements in both the hybrid catalogue and S3M. It is evident that the distributions are essentially identical.

As a further check, we compared MPCORB to the objects that were removed from S3M, since these should have nearly identical distributions. In Figure A1, we show a comparison of the densities for the heliocentric $x$ and $y$ and it is clear that these distributions are left unchanged in the hybrid catalogue.

### 2.2. *Simulated Observations*

In order to investigate the effect of LSST sources on the NEOCP, we use simulated observations of the hybrid catalogue.

[TODO: Should probably talk about whatever Sam did here (needs citation)]

### 2.3. `digest2` *Score Calculation*

The main criterion for an object to be placed on the NEOCP is for it to receive an NEO score of at least 65. This score from 0 to 100 assesses the probability that the object is an NEO and is calculated using the `digest2` code (Keys et al. 2019). We therefore use the same code to check which of the simulated observations of the hybrid catalogue could be submitted to the NEOCP.

We use `digest2` to calculate the NEO score of each NEO and MBA in the hybrid catalogue that is observed in the first year. We apply three further cuts before considering which submissions are eligible for the NEOCP.

1. **Number of observations:** We consider only objects which have at least 2 observations on a given night (though we investigate the effect of making this limit more stringent below)

2. **Minimum arc length:** We ensure that each arc is at least 1 arcsecond in length (needs citation)

3. **Maximum time separation:** We set the maximum time between observations to 90 minutes (needs citation). Thus we only allow tracklets that have at least one pair of observations that occur within 90 minutes of each other

With these cuts and the NEO scores of each object we are able to assess the load on the NEOCP from LSST.

## 3. RESULTS

### 3.1. *Traffic and Purity of the NEOCP*

In Figure 2, we summarise the effect of LSST submissions on the NEOCP. The top panel shows the traffic of the NEOCP, meaning the number of objects that would be submitted to the page, whilst the bottom panel shows the purity, meaning the fraction of objects submitted to the page that are actually NEOs. This plot shows the result of submitting *every* qualifying observation to the NEOCP.

The current typical traffic of the NEOCP is on the order of 25 submissions per night (needs citation). We show in the top panel that this traffic would increase by up to 3 orders of magnitude as a result of LSST submissions. Each line corresponding to a different number of minimum nightly observations (increasing from the original choice of 2 as discussed in Section 2.3). Although the traffic is lower when requiring more observations, even with a minimum of 6 observations the traffic can reach several thousands of submissions per night, which is far more than the NEOCP is currently equipped to handle.

The purity of the NEOCP is also severely impacted by LSST observations, with the abundance of MBA observations polluting the page as false NEOs. For almost the entire year the page will have a purity below 10%, meaning that only 1 in 10 objects on the page is actually an NEO. This can decrease below even 1% and thus a lot of follow up time would be wasted in looking at MBAs masquerading as NEOs.

There are two periodic effects that can be noted in both the traffic and the purity panels. There is a clear seasonal variation over the year as the ecliptic plane moves through the sky. LSST observes from the southern hemisphere and thus in first half of the year when the ecliptic is at lower declinations, more MBAs will be observed. This both increases the traffic and decreases the purity of the NEOCP. In addition to this long term variation, there are also shorter term variations that are most clearly seen in the purity panel. On day 17 and approximately every 30 days after this there is a periodic decrease in the purity of the page. This coincides with the occurrence of a full moon, since LSST has to observe a different part of the sky during this time and this leads to more MBA detections and thus lower purity.

[TODO: Check that I explained both of these effects correctly (particularly the moon one)]

Overall, it is clear that if we proceed in the same manner as is currently recommended that the NEOCP will
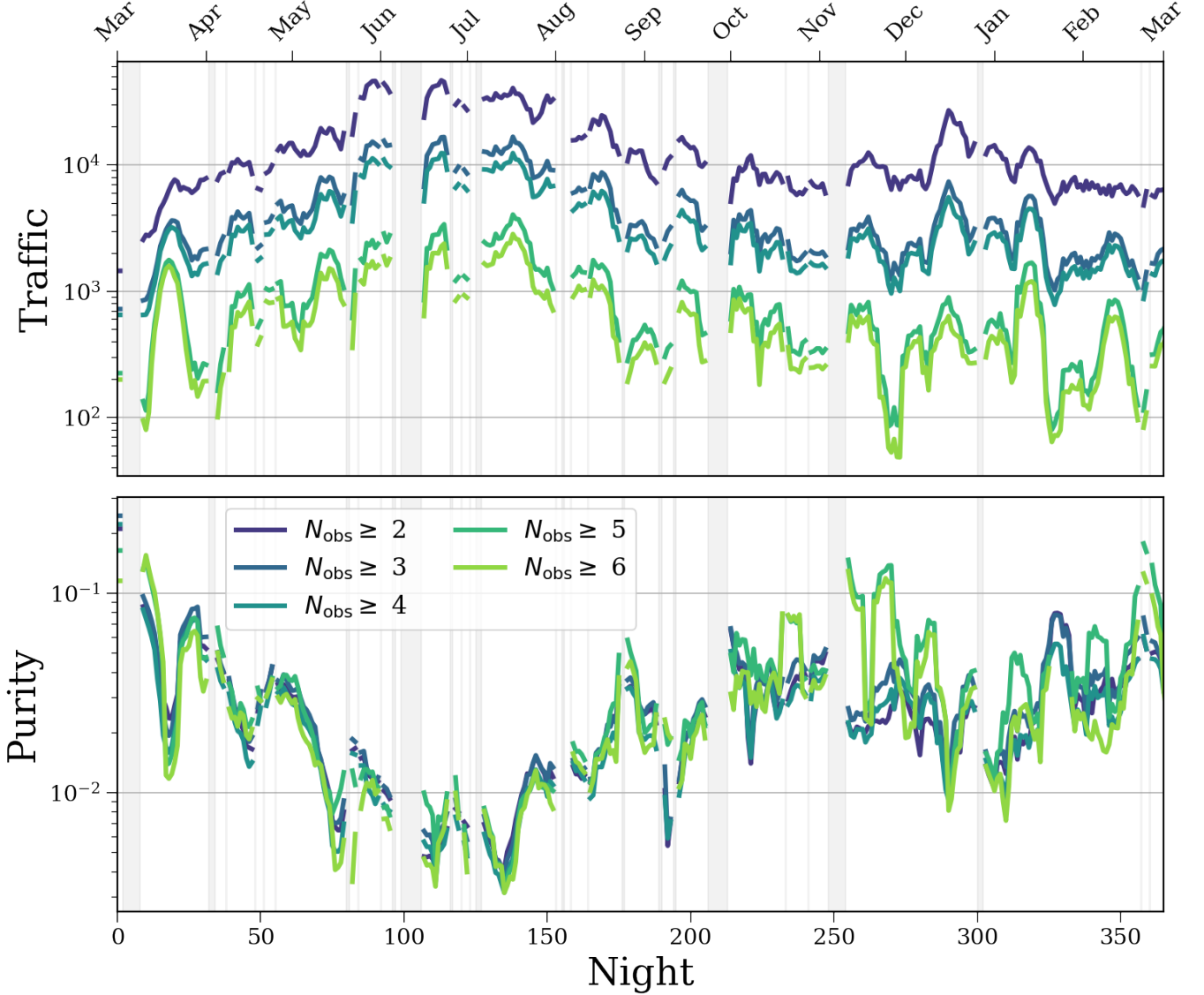
**Figure 2.** Traffic (number of objects sent) and purity (fraction of objects sent that are NEOs) of the NEOCP during the first year of LSST if every observation that qualifies for submission is submitted. Each line is plotted using a rolling window of a week to smooth stochastic effects. Different lines correspond to different constraints on the number of observations for a tracklet to be submitted. Nights on which no observations were taken are highlighted with grey areas.

not be able to handle the load. We now consider how we could proceed differently.

### 3.2. *NEOCP without LSST Detections*

LSST will be able to detect and characterise many potential NEOs without any external input. This means that many of the objects submitted to the NEOCP will actual result in wasted follow up time by the community. We assess the magnitude of this effect using the python package `difi`[3] (needs citation). This package

calculates which objects are detected by LSST, where a detection is defined as occurring when the same object is observed on at least 3 separate nights, each with at least 2 observations separated by at most 90 minutes (needs citation).

In Figure 3, we repeat Figure 2 but now only including objects that would not be detected by LSST alone. This represents a best case scenario in which we had foreknowledge of which observed objects would be detected. In this scenario we see that the traffic from LSST rarely exceeds more than 100 objects and is often only a handful. Similarly the purity is much higher, with an average of around 25%. We conclude that if we were to know
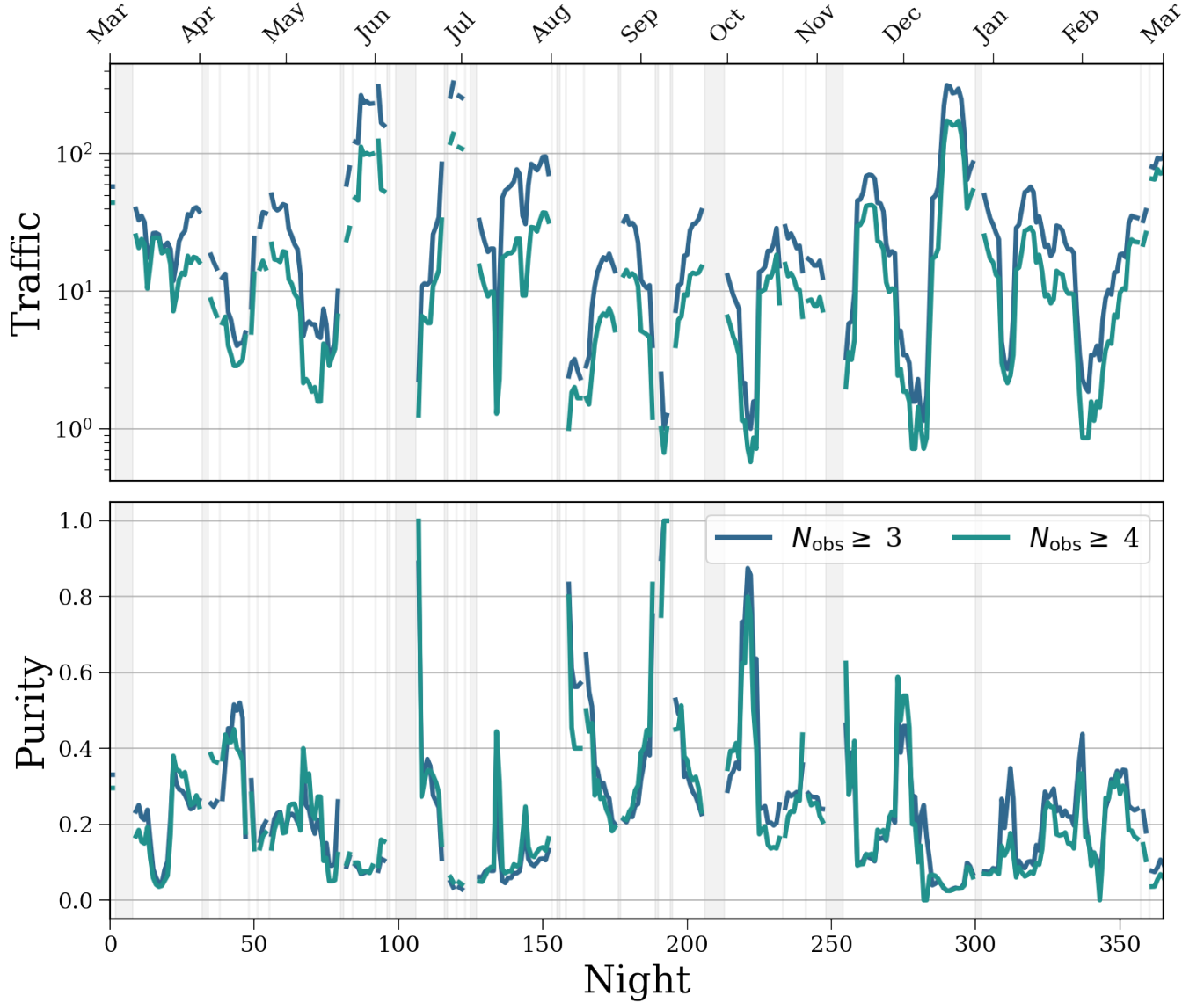
---

[3]https://github.com/moeyensj/difi

**Figure 3.** As Figure 2, but only including objects that would not be detected by LSST alone. Note the y-scale of the lower panel is now linear.

whether an object would be detected by LSST, the load on the NEOCP would remain at manageable levels.

### 3.3. *Estimating LSST Detection Probability*

We now propose a method of *predicting* whether an observed object will be detected by LSST in order to emulate the best case scenario described in Section 3.2.

For an object to be detected by LSST, it must be observed at least twice on at least three separate nights. Therefore, one can predict whether an object will be detected simply by whether it will be in LSST's field of view on the following night - if it will pass out of the field of view by the next night then it will not be detected by LSST and should be submitted to the NEOCP.

In order to determine the location of an object on the following night one needs to know its orbit. Each object on a given night consists of at least two observations

$$O = \{\alpha, \delta, t\} \tag{1}$$

where $\alpha$ is the right ascension, $\delta$ is the declination and $t$ is the time of observation. One can then determine the proper motion of the object on the sky, $(\dot{\alpha}, \dot{\delta})$, by calculating the change in position over time between the two observations. This determines 4 and the 6 orbital elements, but the distance, $D$ and radial velocity, $\dot{D}$ of the object are unconstrained.

We draw a sample of $D$ uniformly in log-space between 0.1-100 AU and a sample of $\dot{D}$ uniformly between

$-50$-$50\,\mathrm{km\,s^{-1}}$ to create a grid of $(D, \dot{D})$. Combining these with the measured $(\alpha, \delta, \dot{\alpha}, \dot{\delta})$ values (with some scatter to account for the detector uncertainty), we create a series of possible variant orbits for the object. We adjust these orbits to account for the light travel time using THOR (Moeyens et al. 2021) and then propagate them until the next night using OpenOrb (Granvik et al. 2009). [TODO: Check through this paragraph to check we actually do this once I actually do it haha].

We estimate the probability of the object being detected by LSST, $P(\text{detect})$, as

$$P(\text{detect}) \propto \sum_{i=1}^{N} \sum_{j=1}^{M} \theta_{i,j} P(D_i) P(\dot{D}_j) \qquad (2)$$

where $\theta_{i,j}$ is a heaviside step function that is 1 if the orbit corresponding to distance $D_i$ and radial velocity $\dot{D}_j$ results in the object remaining in the LSST field of view and 0 otherwise, $P(D_i)$ is the probability of the distance $D_i$ occurring in the hybrid catalogue and $P(\dot{D}_j)$ is the probability of the radial velocity $\dot{D}_j$ occurring in the hybrid catalogue.

[TODO: Need to explain how we know where the detector is]

[TODO: Demonstrate the efficiency of this method with some plots]

[TODO: Maybe also add plots of the D/dot(D) distributions of HC]

## 4. DISCUSSION

[TODO: Some of the results may move here, unsure what else]

## 5. CONCLUSION & SUMMARY

[TODO: ]

¹ Acknowledgements

*Software:* digest2 v0.19.2 (Keys et al. 2019), OpenOrb (Granvik et al. 2009), difi, THOR (Moeyens et al. 2021), astroML (VanderPlas et al. 2012, 2014), scipy (Virtanen et al. 2020)

## REFERENCES

Granvik, M., Virtanen, J., Oszkiewicz, D., & Muinonen, K. 2009, M&PS, 44, 1853, doi: 10.1111/j.1945-5100.2009.tb01994.x

Grav, T., Jedicke, R., Denneau, L., et al. 2011, PASP, 123, 423, doi: 10.1086/659833

Keys, S., Vereš, P., Payne, M. J., et al. 2019, PASP, 131, 064501, doi: 10.1088/1538-3873/ab1157

Moeyens, J., Jurić, M., Ford, J., et al. 2021, AJ, 162, 143, doi: 10.3847/1538-3881/ac042b

VanderPlas, J., Connolly, A. J., Ivezic, Z., & Gray, A. 2012, in Proceedings of Conference on Intelligent Data Understanding (CIDU, 47–54, doi: 10.1109/CIDU.2012.6382200

VanderPlas, J., Fouesneau, M., & Taylor, J. 2014, AstroML: Machine learning and data mining in astronomy, Astrophysics Source Code Library, record ascl:1407.018. http://ascl.net/1407.018

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261, doi: 10.1038/s41592-019-0686-2
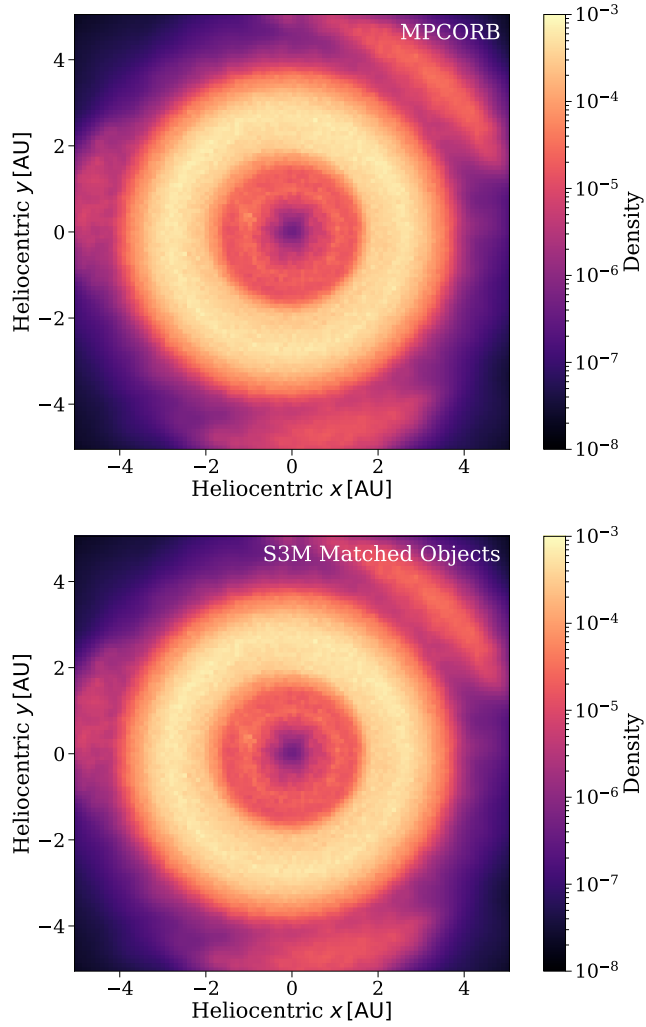
APPENDIX

## A. EXTRA FIGURES





**Figure A2.**

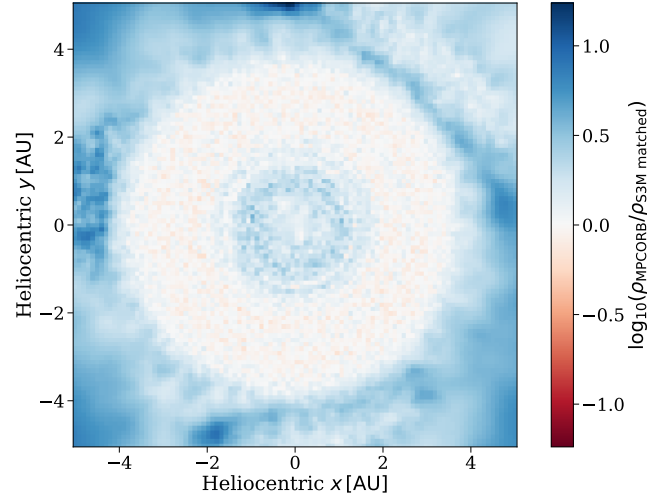**Figure A1.** A comparison of the density of MPCORB objects with those objects that were matched in S3M by our hybrid catalogue pipeline.