

# The Sky is Falling?

## The NEOCP in the Era of LSST

T. WAGG <sup>1</sup>, M. JURIC <sup>1</sup> AND . . . MORE

<sup>1</sup>*Department of Astronomy, University of Washington, Seattle, WA, 98195*

(Received September 6, 2022)

### ABSTRACT

**[TODO: ]** Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur 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.

**Keywords:** Near-Earth objects, Asteroids, Solar system, Small Solar System bodies, Surveys

### 1. INTRODUCTION

Near-Earth Objects (NEOs) are objects that may come close enough to the Earth that small perturbations in their orbit lead to a collision course. A simple definition for NEOs is that there are small solar system bodies that have a perihelion less than 1.3 AU (**needs citation**). Given the threat posed by these objects, it is therefore critical that we catalogue and determine the orbit of NEOs.

The Minor Planet Center maintains a catalogue of known NEOs and their orbits<sup>1</sup>, as well as the NEO confirmation page (NEOCP<sup>2</sup>), which lists potential NEOs that should be prioritised for additional observations by

the NEO follow-up community. An object is only listed on the NEOCP when it has a high probability of being an NEO. This probability is quantified using the **digest2** code, which assigns a score between 0 and 100 and only objects with a score of 65 or more are listed on the page (Keys et al. 2019). Currently, on average around two dozen objects are added to the NEOCP on each night.

The Legacy Survey of Space and Time (LSST) will rapidly increase the rate at which NEOs are discovered and sent to the NEOCP. **[TODO: need more details]**

Given this increased discovery rate, if LSST operates in the same way as current telescopes that work on NEO discovery, the number of objects submitted to the NEOCP will increase by several orders of magnitude. This will make it difficult for the follow-up community to prioritise particular objects and may lead to some objects being missed. However, many of the objects that LSST observes will be re-observed and discovered by LSST alone, though there will still be a sizeable frac-

Corresponding author: Tom Wagg  
tomjwagg@gmail.com

<sup>1</sup><https://www.minorplanetcenter.net/iau/MPCORB/NEA.txt>

<sup>2</sup><https://www.minorplanetcenter.net/iau/NEO/toconfirm-tabular.html>

tion of objects that require community follow-up. It is therefore important to identify this fraction so that the follow-up community can effectively prioritise observations.

This paper quantifies the impact of LSST on the NEOCP and the fraction of objects that require community follow-up.

Our paper is structured as follows. In Section 2, we describe the methods for creating a hybrid catalogue of objects, simulating LSST observations and calculating `digest2` scores for the NEOCP. [TODO: Outline rest of paper]

## 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 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<sup>3</sup>.

#### 2.1.1. Data preprocessing

For the synthetic catalogue of the solar system with we use S3M, the Pan-STARRS Synthetic Solar System Model (S3M Grav et al. 2011). We merge this synthetic catalogue with the latest version of MPCORB<sup>4</sup>, 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. We assess the quality of the hybrid catalogue in Appendix A.

### 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)

<sup>3</sup>[https://github.com/dirac-institute/hybrid\\_sso\\_catalogue/tree/main/hybridcat/hybridcat](https://github.com/dirac-institute/hybrid_sso_catalogue/tree/main/hybridcat/hybridcat)

<sup>4</sup><https://minorplanetcenter.net//iau/MPCORB.html>

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 1, 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 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` (Moeyens 2021). 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 within a 15 day window, each with at least 2 observations separated by at most 90 minutes (needs citation).

In Figure 2, we repeat Figure 1, 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 whether an object would be detected by LSST, the load on the NEOCP would remain at manageable levels.

## 4. DISCUSSION

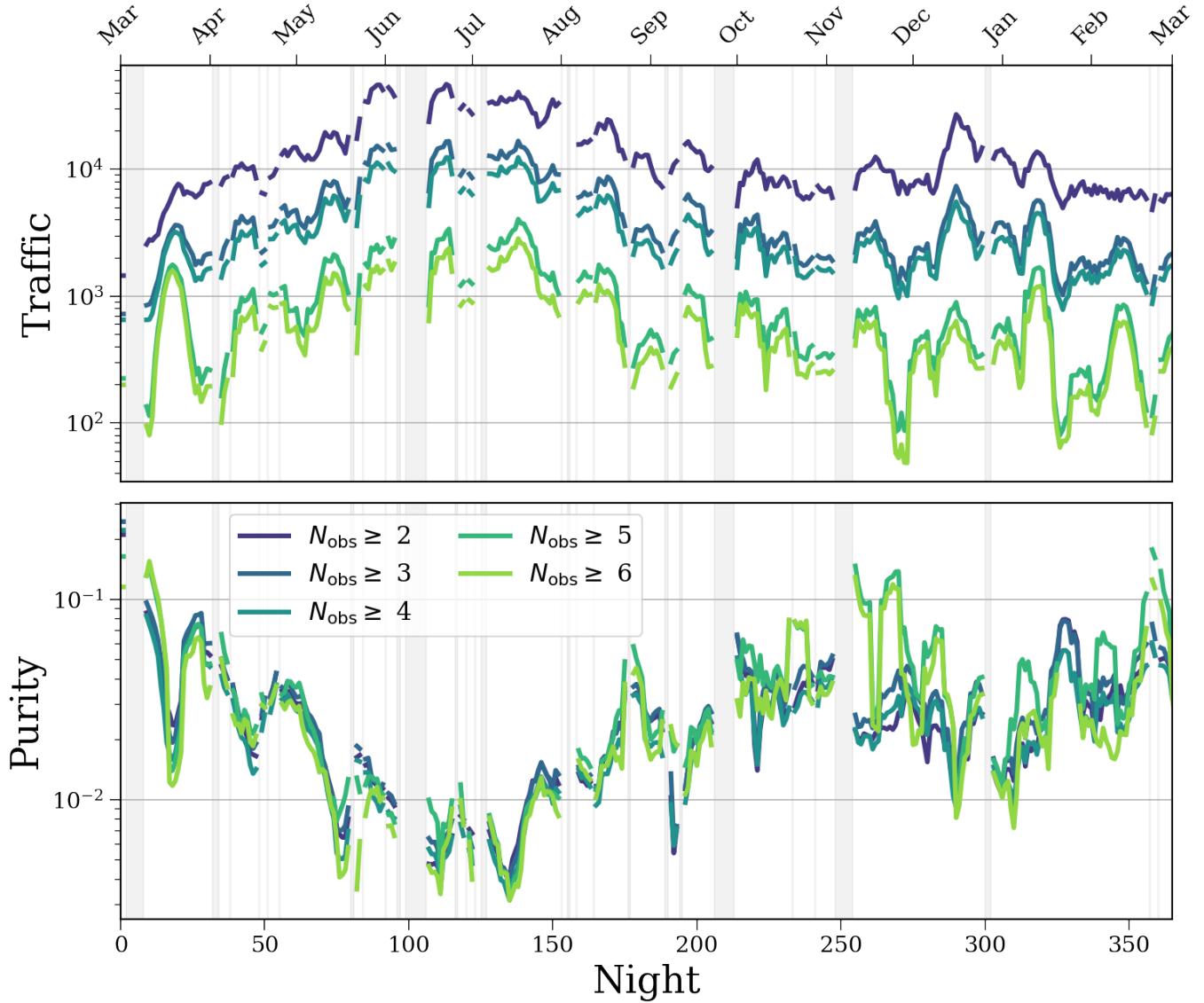
#### 4.1. Estimating LSST Detection Probability

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. This requires one to predict both the location and magnitude of the object on subsequent nights, as well as the visit schedule of LSST.

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

$$O = \{\alpha, \delta, t\} \quad (1)$$

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



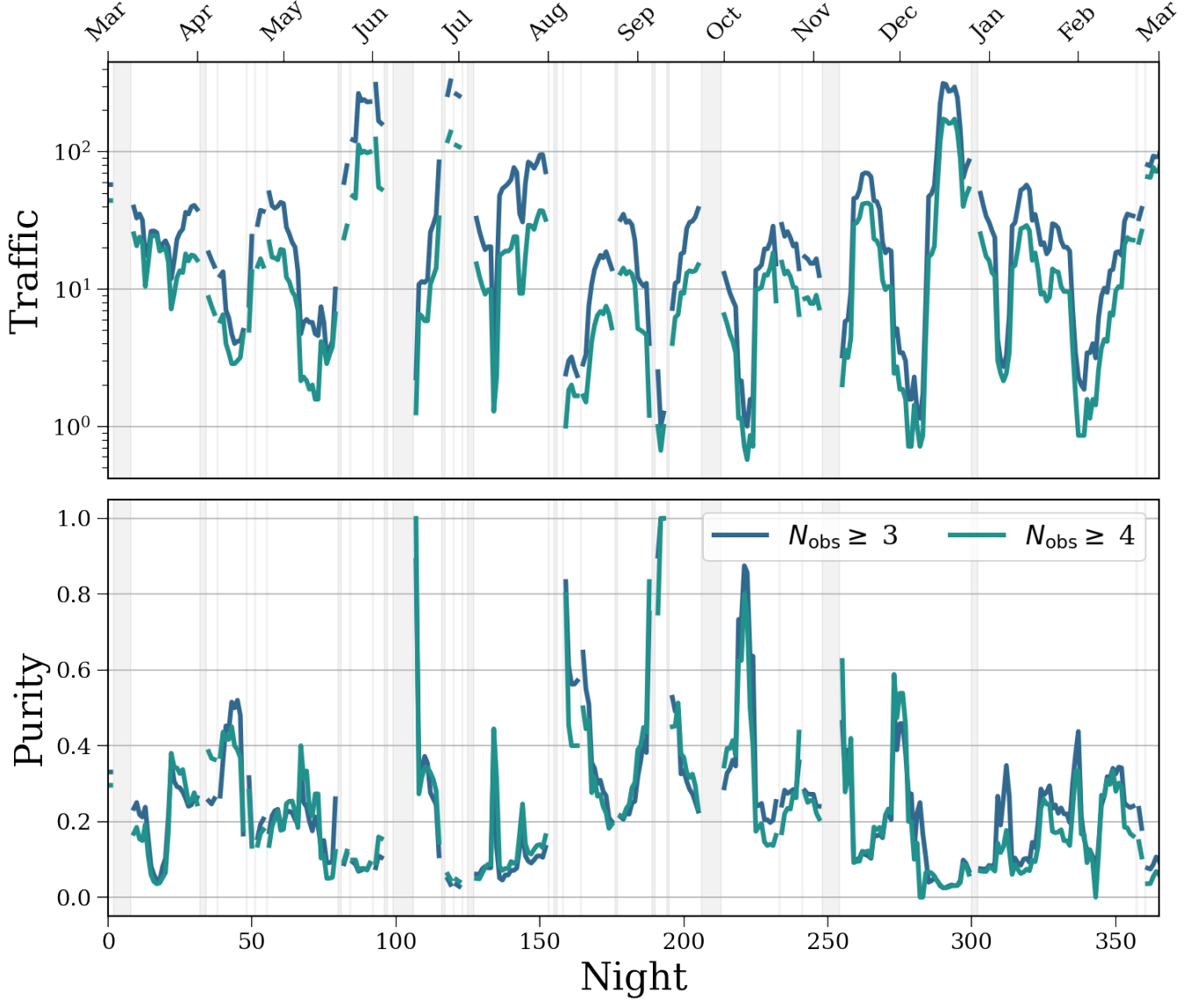
**Figure 1.** Traffic (number of objects sent) and purity (fraction of objects sent that are NEOs) of the NEOCP during the first year of LSST if every observation that qualifies for submission is submitted. Each line is plotted using a rolling window of a week to smooth stochastic effects. Different 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.

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  $[-100, 100]$  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 and adjust them to account for the light travel time using THOR (Moeyens et al. 2021). Since we are only interested in NEOs, we mask out any orbits that have a perihelion distance of greater than 1.3 AU. Additionally,

we determine the absolute magnitude of the object for each orbit using the HG-system. We convert the apparent magnitudes of the observations in the first night to V-band magnitudes using the LSST colour definitions, assuming a C-type asteroid (Jones et al. 2018). We then use the mean observed V-band magnitude, the assumed distances and fixed slope parameter of  $G = 0.15$  (needs citation) to calculate the current absolute magnitude.

Finally, using `OpenOrb`, we produce ephemerides for the orbits for each visit in the following 14 nights of the predicted schedule (Granvik et al. 2009). For each orbit and each field, we assume an observation will occur if the object is within  $2.1^\circ$  of the centre of the field and has



**Figure 2.** As Figure 1, but only including objects that would not be detected by LSST alone. Note the y-scale of the lower panel is now linear.

an apparent magnitude brighter than the  $5\sigma$  depth of field. Each orbit is defined as producing a detection if it has at least 2 observations on at least 3 nights in the 15 day window, requiring that the minimum arc length is 1 arcsecond and that the maximum separation between observations is 90 minutes.

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

$$P(\text{detect}) \propto \sum_{i=1}^{N_{\text{orbits}}} \theta_i w_i \quad (2)$$

where  $\theta_i$  is a heaviside step function that is 1 if the orbit produces a detection and 0 otherwise, and  $w_i$  is

the weight associated with that orbit [TODO: which is yet to be determined].

#### 4.2. Accuracy of prediction algorithm

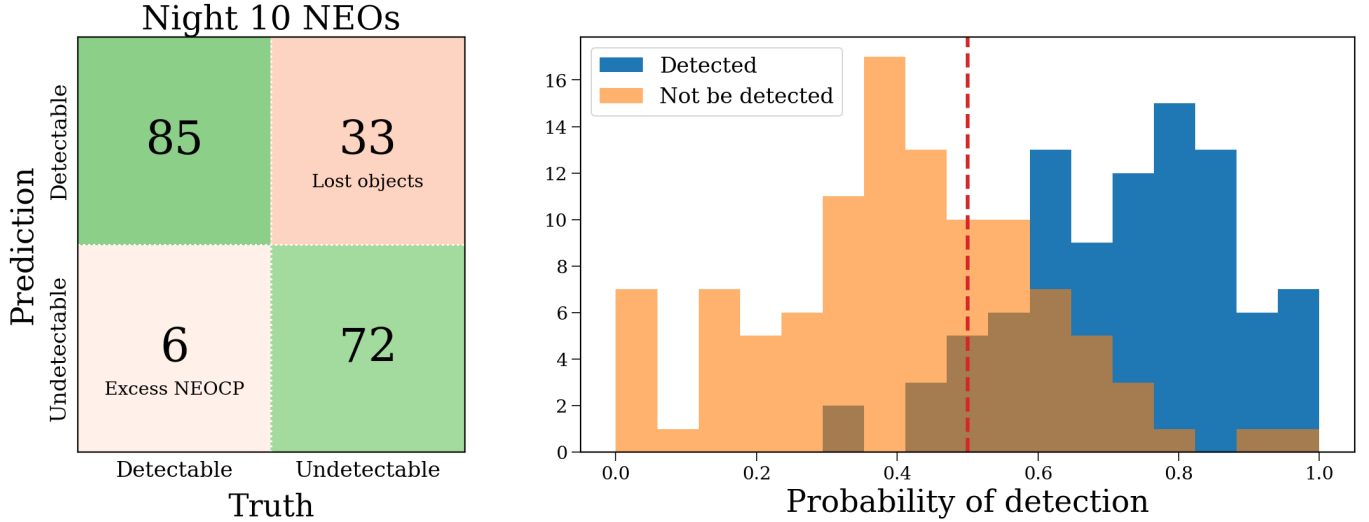
[TODO: Demonstrate the efficiency of this method with some plots like fig 3]

### 5. CONCLUSION & SUMMARY

[TODO: Write and summary and spruce up the items below]

1. **Hybrid Catalogue:** We produce and make available a hybrid catalogue that combines a simulated catalogue (Grav et al. 2011) with detections from MPCORB. This distributions of the six orbital elements as well as the absolute magnitude in





**Figure 3.** A demonstration of the prediction algorithm described in Section 4.1. **Left:** A confusion matrix representing the algorithm’s ability to predict the detectability of an object. **Right:** A distribution of assigned probabilities, split into a population of objects that are actually detected and a population that are not.

this catalogue are unchanged from S3M. The hybrid catalogue can be used to make predictions for LSST that account for prior detections (see Section 2.1).

2. **NEOCP traffic and purity from LSST:** We show that the traffic (number of objects submitted) and purity (fraction of objects that are NEOs) of the NEOCP will be severely impacted by LSST. We demonstrate that if LSST follows current submission guidelines that the NEOCP would be unable to handle the volume of submissions

3. **NEOCP without LSST detections:** We highlight that the traffic and purity of the NEOCP remain at acceptable levels if LSST only submits objects that it won’t detect alone

4. **LSST detection probability prediction:**

## 1 Acknowledgements

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

## REFERENCES

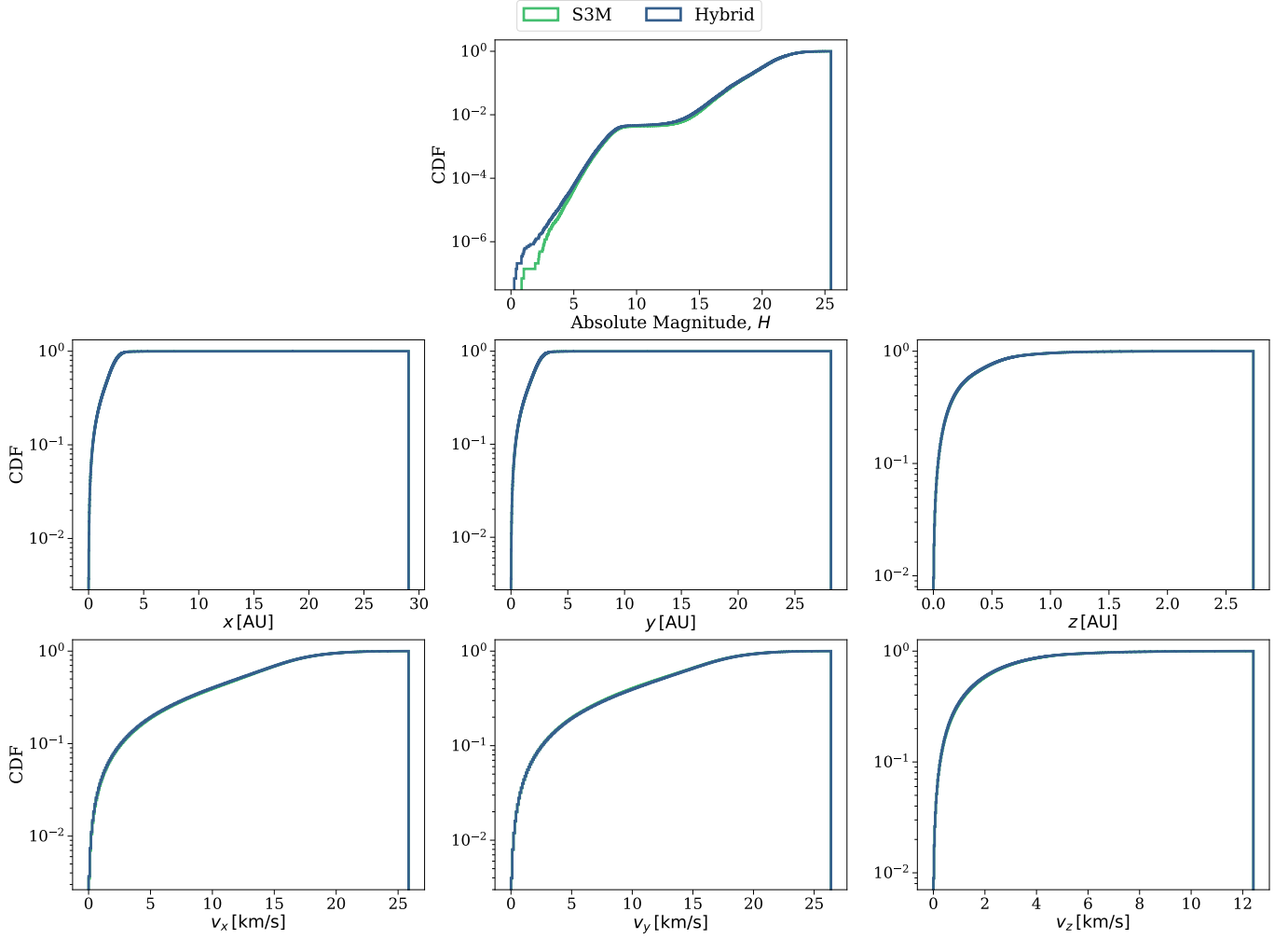
- Granvik, M., Virtanen, J., Oszkiewicz, D., & Muinonen, K. 2009, *M&PS*, 44, 1853, doi: [10.1111/j.1945-5100.2009.tb01994.x](https://doi.org/10.1111/j.1945-5100.2009.tb01994.x)
- Grav, T., Jedicke, R., Denneau, L., et al. 2011, *PASP*, 123, 423, doi: [10.1086/659833](https://doi.org/10.1086/659833)
- Jones, R. L., Slater, C. T., Moeyens, J., et al. 2018, *Icarus*, 303, 181, doi: [10.1016/j.icarus.2017.11.033](https://doi.org/10.1016/j.icarus.2017.11.033)
- Keys, S., Vereš, P., Payne, M. J., et al. 2019, *PASP*, 131, 064501, doi: [10.1088/1538-3873/ab1157](https://doi.org/10.1088/1538-3873/ab1157)
- Moeyens, J. 2021, moeyensj/difi: v1.1, v1.1, Zenodo, doi: [10.5281/zenodo.4752719](https://doi.org/10.5281/zenodo.4752719)
- Moeyens, J., Jurić, M., Ford, J., et al. 2021, *AJ*, 162, 143, doi: [10.3847/1538-3881/ac042b](https://doi.org/10.3847/1538-3881/ac042b)
- VanderPlas, J., Connolly, A. J., Ivezić, Z., & Gray, A. 2012, in *Proceedings of Conference on Intelligent Data Understanding (CIDU)*, 47–54, doi: [10.1109/CIDU.2012.6382200](https://doi.org/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](https://doi.org/10.1038/s41592-019-0686-2)

## APPENDIX

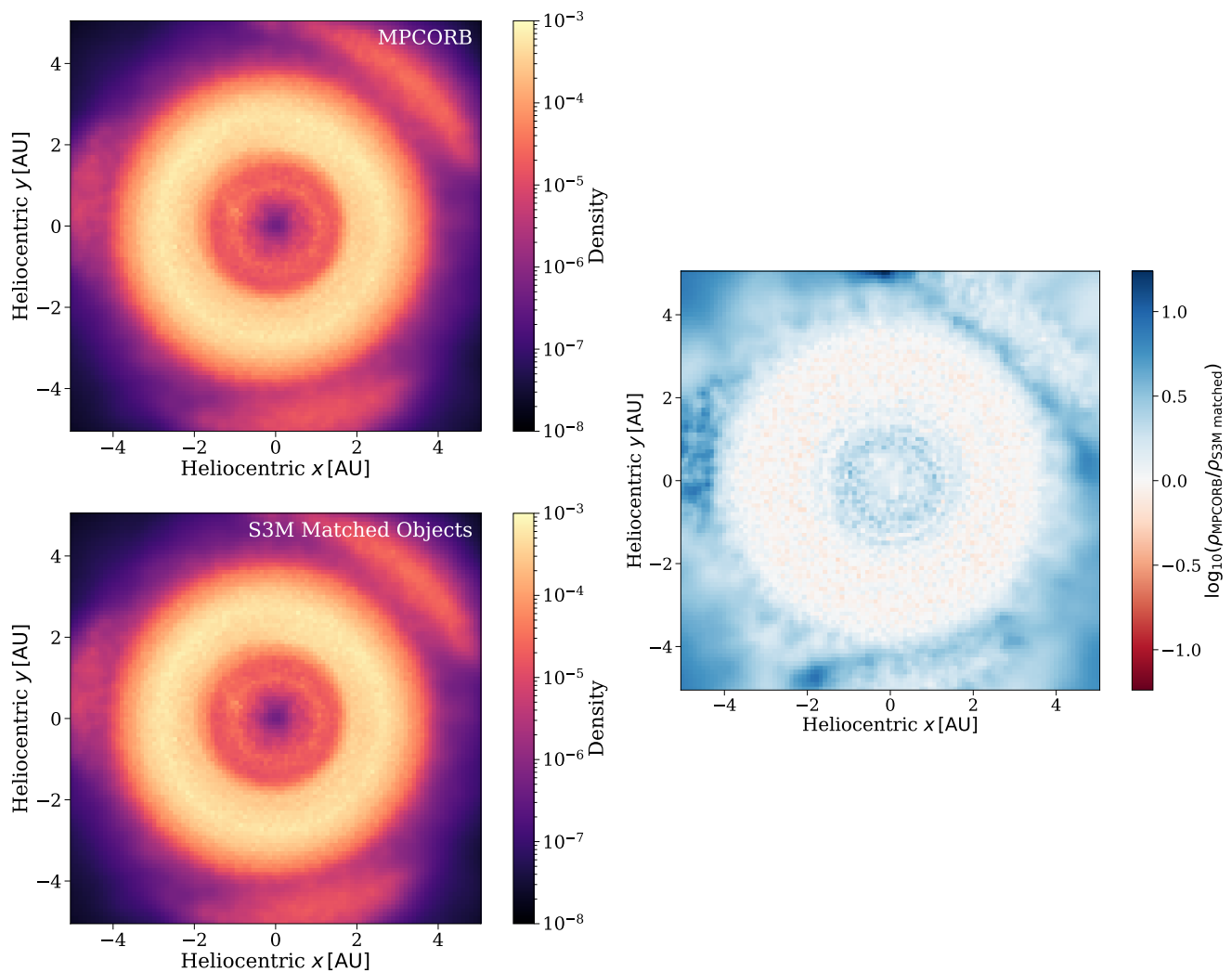
## A. ASSESSING QUALITY OF HYBRID CATALOGUE

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 A1, 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 A2, 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.



**Figure A1.** A comparison of the parameter distributions of S3M (Grav et al. 2011) and the hybrid catalogue we created.



**Figure A2.** **Left:** A comparison of the density of MPCORB objects with those objects that were matched in S3M by our hybrid catalogue pipeline. **Right:** Residuals between the two plots in the left panel.