The DECam Ecliptic Exploration Project (DEEP). III. Survey Characterization and Simulation Methods

Pedro H. Bernardinelli[1,16], Hayden Smotherman[1], Zachary Langford[1,2], Stephen K. N. Portillo[1,3], Andrew J. Connolly[1],
J. Bryce Kalmbach[1], Steven Stetzler[1], Mario Jurić[1], William J. Oldroyd[4], Hsing Wen Lin (林省文)[5],
Fred C. Adams[5,6], Colin Orion Chandler[1,4,7], Cesar Fuentes[8], David W. Gerdes[5,6], Matthew J. Holman[9],
Larissa Markwardt[5], Andrew McNeill[4,10], Michael Mommert[11], Kevin J. Napier[5], Matthew J. Payne[9],
Darin Ragozzine[12], Andrew S. Rivkin[13], Hilke Schlichting[14], Scott S. Sheppard[15], Ryder Strauss[4],
David E. Trilling[4], and Chadwick A. Trujillo[4]

[1] DiRAC Institute and the Department of Astronomy, University of Washington, Seattle, WA, USA; phbern@uw.edu
[2] Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA, USA
[3] Department of Mathematical and Physical Sciences, Concordia University of Edmonton, 7128 Ada Boulevard, Edmonton, AB, T5B 4E4, Canada
[4] Department of Astronomy and Planetary Science, Northern Arizona University, P.O. Box 6010, Flagstaff, AZ 86011, USA
[5] Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA
[6] Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA
[7] LSST Interdisciplinary Network for Collaboration and Computing, 933 N. Cherry Avenue, Tucson, AZ 85721, USA
[8] Departamento de Astronomía, Universidad de Chile, Camino del Observatorio 1515, Las Condes, Santiago, Chile
[9] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, MS 51, Cambridge, MA 02138, USA
[10] Department of Physics, Lehigh University, 16 Memorial Drive East, Bethlehem, PA 18015, USA
[11] School of Computer Science, University of St. Gallen, Rosenbergstrasse 30, CH-9000 St. Gallen, Switzerland
[12] Department of Physics and Astronomy, Brigham Young University, Provo, UT 84602, USA
[13] Applied Physics Lab, Johns Hopkins University, 11100 Johns Hopkins Road, Laurel, MD 20723, USA
[14] Department of Earth, Planetary and Space Sciences, University of California Los Angeles, 595 Charles E. Young Drive East, Los Angeles, CA 90095, USA
[15] Earth and Planets Laboratory, Carnegie Institution for Science, Washington, DC 20015, USA

Abstract

We present a detailed study of the observational biases of the DECam Ecliptic Exploration Project's B1 data release and survey simulation software that enables direct statistical comparisons between models and our data. We inject a synthetic population of objects into the images, and then subsequently recover them in the same processing as our real detections. This enables us to characterize the survey's completeness as a function of apparent magnitudes and on-sky rates of motion. We study the statistically optimal functional form for the magnitude, and develop a methodology that can estimate the magnitude and rate efficiencies for all survey's pointing groups simultaneously. We have determined that our peak completeness is on average 80% in each pointing group, and our magnitude drops to 25% of this value at $m_{25} = 26.22$. We describe the freely available survey simulation software and its methodology. We conclude by using it to infer that our effective search area for objects at 40 au is 14.8 deg^2, and that our lack of dynamically cold distant objects means that there at most 8×10^3 objects with $60 < a < 80$ au and absolute magnitudes $H \leqslant 8$.

Unified Astronomy Thesaurus concepts: Kuiper Belt (893); Trans-Neptunian objects (1705); Surveys (1671)

1. Introduction

Our understanding of the trans-Neptunian region has improved significantly since the discovery of the first trans-Neptunian object (TNO) by Jewitt & Luu (1993). Thousands of objects have now been discovered, individual surveys are now capable of finding hundreds of objects (e.g., Petit et al. 2011; Bannister et al. 2016, 2018; Bernardinelli et al. 2020a, 2022), and models of this region make quantitative predictions of the orbital structure, number of objects, and size distribution of this population (see reviews by Morbidelli & Nesvorný 2020; Gladman & Volk 2021). However, such predictions are not immediately testable against the observations from any given survey. Any survey has observational biases, driven primarily by the observational strategy used and the apparent magnitude limits achieved in the exposures. Thus, any comparison of a model to the real trans-Neptunian region (as seen by each survey) must factor in these effects.

Survey simulators have become widespread tools for testing these population models against observations (Jones et al. 2006; Lawler et al. 2018a; Bernardinelli et al. 2022). These tools provide a standardized way to estimate discovery biases of a survey for any given synthetic object as a function of orbital elements, magnitude, colors, and light-curve properties, without requiring the need to directly inject objects into the survey images, which can oftentimes be a prohibitively expensive task, in terms of computation time, especially if multiple models are being tested. A detailed understanding of these biases means that one can not only test the populations of detected objects but also place meaningful limits given survey nondetections in a region of parameter space. Examples in the literature include, but are not limited to, tests for clustering of high semimajor axis (a) and high perihelion (q) trans-Neptunian objects (Shankman et al. 2017; Bernardinelli et al. 2020b; Napier et al. 2021), limits on

---

[16] DiRAC Postdoctoral Fellow.

the discoverability of distant planets in the Solar System (Belyakov et al. 2022), tests of the structure of the classical Kuiper Belt (Petit et al. 2011; Kavelaars et al. 2021; Bernardinelli et al. 2022), and size distribution calculations (Lawler et al. 2018b).

The DECam Ecliptic Exploration Project (DEEP; Trilling et al. 2024) is an ongoing survey using the Dark Energy Camera (DECam; Flaugher et al. 2015) dedicated to finding faint trans-Neptunian objects with digital tracking, using a carefully chosen observational strategy (Trujillo et al. 2024) that maximizes the discovery potential of low-inclination Kuiper Belt objects. The survey consists of a series of "pointing groups" or "long stares," with a nominal observing strategy of 100 images per night in an expanding cone-shaped pattern that follows the typical motion of the TNO population (see Figure 2 of Trujillo et al. 2024). The survey consists of four fields (A0, A1, B0, and B1) that started being imaged in 2019. The analysis presented here uses data from the B1 field with data taken between 2019 and 2021. Applications of the techniques presented here to the other DEEP fields are forthcoming. This observational strategy enables the DEEP project to reach a unique combination of depth and area, which will enable us to probe the absolute magnitude distribution of TNOs as a function of subpopulation with high statistical significance at magnitudes fainter than $m_r \approx 25$, beyond the limits of current (Bannister et al. 2018; Bernardinelli et al. 2022) and upcoming (Ivezić et al. 2019) large area surveys (see Trilling et al. 2024, for a detailed description of the DEEP survey goals).

In Section 2, we present a synthetic population of objects injected into the DEEP images, and in Section 3, we discuss how this population can be used to characterize the survey's completeness as a function of magnitudes and orbital elements. We then use these results to construct a survey simulator, presented in Section 4 and publicly released on `GitHub`.[17] Section 5 presents two quantitative analyses of "toy models" of the trans-Neptunian region to demonstrate the power of the survey simulator, and we conclude in Section 6 by outlining expected improvements.

2. Synthetic Population Injected in the DEEP Data

The DEEP images were processed using the GPU-based `KBMOD` algorithm (Whidden et al. 2019; Smotherman et al. 2021). A detailed presentation of the full object recovery methodology is given in Smotherman et al. (2024; see also Napier et al. 2024). The detection "tracklets" (a pair of sky coordinates derived from the digital tracking procedure projected to the beginning and ending of exposures of each long stare) were subsequently linked into orbits. In Smotherman et al. (2024), we demonstrate that this recovery procedure procedure leads to high-confidence recoveries in both injected fakes and known objects, leading to high-quality orbit fits ($\sigma_a/a \approx 10^{-4}$) and negligible contamination. The digital tracking procedure is conducted on a two-dimensional grid of rates of motion r and a positional angle ϕ, where positive ϕ indicates decreasing ecliptic longitude. In the results presented here, our search was limited to $90 < r < 400$ px day^{-1} (1 DECam pixel is $0\overset{''}{.}263$) and $|\phi| < 45$ deg. As seen in Figure 1, this choice exhausts the parameter space of bound orbits from 20 to 80 au, including the entirety of the Kuiper Belt. We note that objects at slightly higher or lower rates can still be recovered by the search,

as the joint fit procedure of Smotherman et al. (2021) corrects for rates approximately outside the search rate that still make it above the detection threshold if it produces a meaningful (but potentially trailed) stack.

We use the LSST Science Pipelines (Jurić et al. 2017) in order to inject these fakes into the images. Objects were injected at the sky coordinates derived from their orbits at the midpoint of each exposure, and used the realistic PSF model derived for each exposure (Smotherman et al. 2024). These fakes were injected after astrometric and photometric calibration of the images and before the creation of coadded templates and differenced images, guaranteeing that our objects have meaningful PSF fluxes that correspond to their nominal magnitudes at that exposure. This also means that any image-processing artifact stemming from the image coaddition and differencing procedures will impact real and injected objects equally. In other words, we have injected a realistic set of synthetic objects into the data that will be detected by our detection pipeline along with the real objects, thus allowing us to determine the detection efficiency for our survey.

Our goal is to fully understand the limits of what can be detected in single-night stacks and subsequently linked to other nights in the DEEP survey, so this population is not constrained by the expected structure of the trans-Neptunian region, but rather to fully characterize the range of parameters where we can expect to discover objects. To do this, it is important to understand the efficiencies in the regions of the parameter space where objects have been found and also where no objects have been detected, so this population (intentionally) goes beyond the boundaries of the known TNOs. By constructing such a population model, instead of, for example, using a full model of the trans-Neptunian region, we ensure that we are not biasing ourselves toward what other projects have seen, which would impact our ability to make statistical statements about our observations. This population is divided in three parts:

1. A low-eccentricity ($e < 0.3$) population in the $30 < a < 80$ au range. This accounts for the classical Kuiper Belt, as well as moderate a and high-q objects. a and e are uniformly distributed in this range, and the inclination distribution is $\propto \sin(i)$ in the $0° \leqslant i \leqslant 60°$ range, leading to an approximately uniform distribution in i for objects inside the survey footprint. The longitude of ascending node Ω, argument of perihelion ω, and mean anomaly $\mathcal{M}$ are uniformly distributed in the $[0°, 360°]$ range.

2. An "excited" population with a logarithmically distributed in the $20 < a < 2000$ au range, and perihelia uniformly distributed between 10 and 100 au. To avoid creating unphysical orbits, objects where we sampled $a < q$ are changed to $e = 0$. The inclinations are distributed as in the first population, but going up to 90°, and Ω and ω are uniformly distributed. Objects with $a > 150$ au and $q > 30$ au are limited to $|\mathcal{M}| \leqslant 30°$, to ensure that these objects are not generated at large heliocentric distances. This choice of parameters leads to a population that reproduces the range of parameters for the scattering and detached TNO populations.

3. A fully isotropic population that exhausts the parameter space of all bound orbits following the prescription of Bernardinelli et al. (2020a). This population covers the distance range between 25 and 1000 au, with half of the synthetic objects uniformly distributed between $25 < d < 80$ au, and the other half logarithmically distributed between
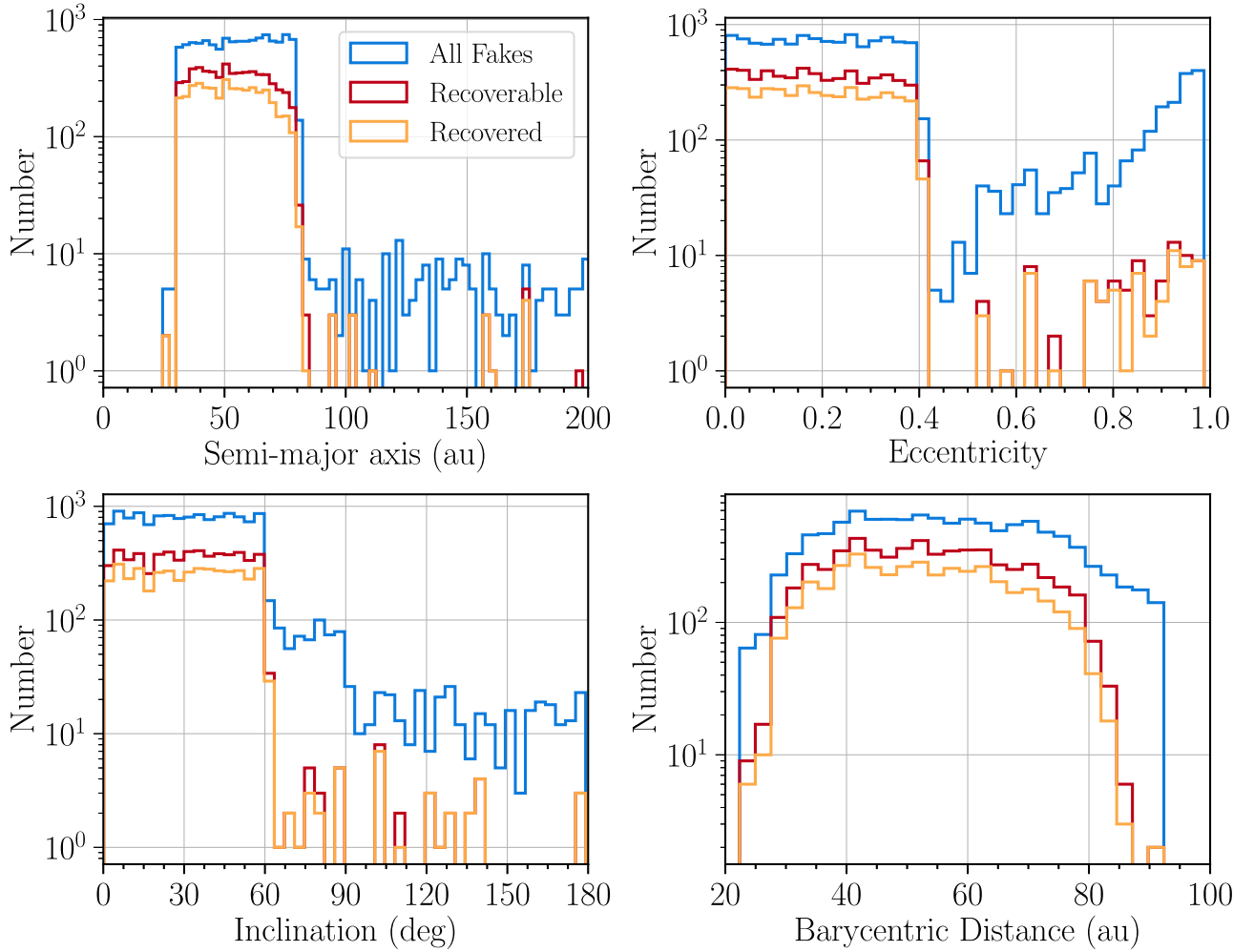
Figure 1. Histograms of semimajor axis a, eccentricity e, inclination i, and barycentric distance d of the injected fakes. The blue lines represent all injected fakes. The red lines ("recoverable" fakes) represent those with rates $150 < r < 400$ px day^{-1} and $m_{VR} < 25.5$, well inside the completeness limit of both rate and magnitude selection functions (as discussed in the Section 3.2). The orange lines indicate those fakes that were recovered by the digital tracking procedure. The histograms of a and d were constrained to the region of parameter space where the majority of the objects are found.

$80 < d < 1000$ au. This population includes objects such as putative distant planets (e.g., Batygin et al. 2019) as well as comets and other highly eccentric orbits at Kuiper Belt distances (e.g., Bernardinelli et al. 2021).

This population is illustrated in Figure 1. We note that each of the histograms of "recoverable" objects (limited to $150 < r < 400$ px day^{-1} and $m_{VR} < 25.5$) and "recovered" objects qualitatively follow each other closely (up to a normalization), and there is no obvious structure in these distributions, i.e., there are no significant dependencies in e and i, while the dependencies in a and d are directly related to the dependency in r (quantified in Section 3). This indicates that our selection function is not dominated by the orbital elements of each object, but rather by the rates and magnitudes. To see that this must be the case, we note that, during the four hours of our long stares, the motion of any given TNO is linear and dominated by the reflex motion of the Earth, scaled by the topocentric distance of the object, suppressing the motion driven by the instantaneous barycentric velocity of this object. In Section 3, we discuss how these parameters can be characterized.

These objects intentionally span a wide range of apparent magnitudes (20% of the sample uniformly distributed between $20 < m_{VR} < 24$ and 80% between $24 < m_{VR} < 28$), significantly beyond the detection limits of the survey, so we can have a complete understanding of the efficiencies and potential pitfalls in the analyses. These magnitudes are assigned to the object for 2020 January 1, and their magnitudes reproduce the expectation given their motion and the motion of the Earth. Half of the objects are also assigned a sinusoidal light curve that changes their magnitude by $\delta m = A \sin(2\pi t/T + \phi_0)$, where the semi-amplitude A is distributed between 0 and 0.5 mag, peaking at $A = 0.25$ mag. The periods T are uniformly distributed between 2 and 100 hr, with a random phases at the 2020 January 1 epoch; we note that, for any period significantly longer than the four hours of the long stare, this means that we are sampling the object at a fully randomized phase for each long stare. This simplified model enables a reasonable characterization of the effect of the light curves in the discoverability of our synthetic objects, and it reproduces the range seen by Strauss et al. (2024) of periods and amplitudes. That is, this choice of magnitude amplitudes broadens our efficiency functions in a realistic manner (so that fainter objects are found if they are at the peak of their light curves, and conversely, brighter objects are missed if their light curve is at a minimum). All reported magnitudes correspond to the mean magnitude of each object

during the corresponding long stare. We will defer the full characterization of our light-curve amplitude selection function to a future publication (see also Strauss et al. 2024).

3. Estimating the Detection Efficiency

The detectability of an object combines both the information on magnitude m and orbital elements $\boldsymbol{a} = \{a, e, i, \Omega, \omega, \mathcal{M}\}$ by factoring in the detection probabilities $p(m)$ with the survey geometry and digital tracking efficiencies $p(\boldsymbol{a}, m|\text{survey geometry}, \text{digital tracking})$, and for objects linked in multiple nights, the linking probability $p(\boldsymbol{a}, \{m_\mu\}|\text{linking})$, accounting for the probability that the object recovered in multiple pointing groups μ is subsequently linked to an orbit and also that such a linkage represents the proper set of detections of the object.[18] While we can easily estimate the magnitude-independent $p(\boldsymbol{a}|\text{survey geometry})$, the digital tracking probability of detection is both a function of magnitude m and rate of motion r. This is distinct from a survey that detects objects in each image, as TNOs are typically very close to point sources, and so their detectability is dominated by the the limiting magnitude of the exposure for an object inside its effective area. We introduce here a methodology that simultaneously accounts for the magnitude and rate completeness of each pointing group.

3.1. Efficiency as a Function of Magnitude

We begin by estimating the detection efficiency as a function of magnitude for each pointing group. We follow the prescription of Bernardinelli et al. (2020a), where we parameterize the efficiency by a probability of detection $p(m|\boldsymbol{\theta})$ for an object with magnitude m using a set of variables $\boldsymbol{\theta}$. After the synthetic objects have been properly recovered and identified, we index the recovered objects by α and nonrecovered by β. We maximize the likelihood:

$$\mathcal{L} = \prod_\alpha p(m_\alpha|\boldsymbol{\theta}) \prod_\beta [1 - p(m_\beta|\boldsymbol{\theta})]. \tag{1}$$

This likelihood allows us to estimate the parameters of our chosen efficiency function without having to bin our injected samples.

There are many possible choices of functional forms for $p(m|\boldsymbol{\theta})$ (e.g., Gladman et al. 1998; Bernstein et al. 2004; Jones et al. 2006; Bannister et al. 2016; Bernardinelli et al. 2020a). The simplest possible form is a step function, where the only free parameters are the limiting magnitude m_0 and the efficiency for $m < m_0$. Smooth functional forms are also possible, with the logistic function

$$p(m|m_{50}, c, k) = \frac{c}{1 + \exp(k(m - m_{50}))} \tag{2}$$

being one of the simplest possible parameterizations. Here, m_{50} is the magnitude of 50% completeness, $c \in [0, 1]$ the peak efficiency, and k the transition sharpness between the "detected" and "not detected" regimes. These parameters are such that $p(m) = c$ for $m \ll m_{50}$, $p(m_{50}) = c/2$, and $p(m) = 0$ for $m \gg m_{50}$. The limit $k \to \infty$ corresponds to a step function. A similar parameterization finds the magnitude of 25%

efficiency (m_{25}) fitting two logistics[19] with transition sharpness k_1 and k_2,

$$p(m|m_{25}, c, k_1, k_2)$$
$$= \frac{c}{[1 + \exp(k_1(m - m_{25}))][1 + \exp(k_2(m - m_{25}))]}. \tag{3}$$

We also posit a third model, which we fit as a sanity check to verify whether we are overfitting our data, that uses three logistics and has a characteristic magnitude of 12.5% completeness $m_{12.5}$. We note that this methodology extends to other functional forms, not just the ones presented here (see, e.g., Bannister et al. 2016, for a functional form that quantifies lost bright objects).

To decide the optimal functional form, we use the Bayesian information criterion (BIC; Schwarz 1978) test on the maximum likelihood $\mathcal{L}$. For N_{par} model parameters and N_{data} data points, we have that the optimal model minimizes

$$\text{BIC} = N_{\text{par}} \log N_{\text{data}} - 2 \log \mathcal{L}. \tag{4}$$

Because we are interested in the general behavior of our efficiency function, we perform this by fitting one $p(m)$ to the entire set of synthetic objects, that is, we estimate an approximate selection function for all long stares. We restrict our data to $150 < r < 400 \, \text{px day}^{-1}$, a region where the dominating cause of object loss is from the magnitude selection function. Of the three functional forms for $p(m)$ presented here, Equation (3) has the lowest BIC of $\text{BIC} = 8055.14$, while Equation (2) and the triple exponential have $\text{BIC} = 8071.99$ and $\text{BIC} = 8061.12$, respectively. Following Kass & Raftery (1995), the difference between two BIC values approximates the Bayes factor between these two models, and therefore, among a set of models (model 1, model 2, and model 3), we can estimate the odds that a given model best fits our data. If we call model i the model with the lowest BIC, then the probability $\mathcal{O}$ that model that best fits the data is approximately

$$\mathcal{O}_{i, j \neq i} = \exp[(\text{BIC}_i - \text{BIC}_j)/2]. \tag{5}$$

The odds against Equation (2) fitting the data better than Equation (3) are approximately 4570:1, and the odds against a triple exponential fitting the data better than Equation (3) are approximately 20:1. We have therefore confirmed that the model given by Equation (3) is likely the best model for our application of the three considered here.

However, we stress that the BIC is an approximate metric and that the choice of model among these three models does not qualitatively alter our characterization, as the m_{50} and m_{25} models predict similar drop-off points. Converting the best-fit $m_{25} = 26.22$ from Equation (3) to an m_{50} value, we get a converted m_{50} of 25.96, while fitting Equation (2) directly leads us to $m_{50} = 25.93$. Our results are therefore only weakly sensitive to the choice of $p(m)$. From now on, we will assume $p(m)$ is given by Equation (3), but our results and methodology are independent of this choice.

---

[18] For our purposes, we do not require that *all* detections of a given object in the survey are recovered—we only require the recovery of a subset that specifies the object.

[19] In the literature, this function is sometimes expressed as two hyperbolic tangent functions. This is equivalent to what we show here.
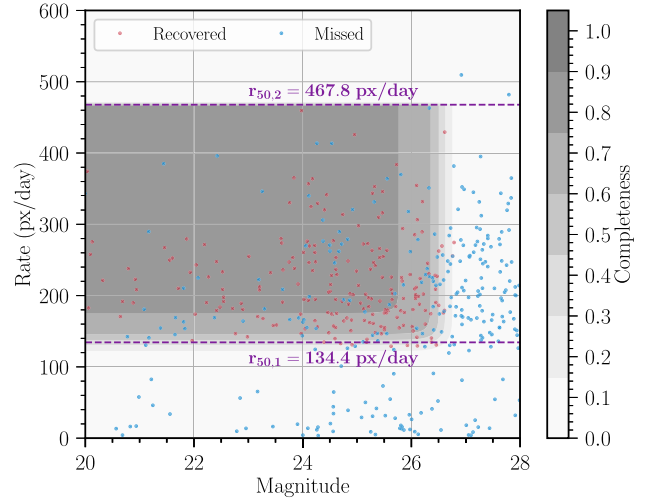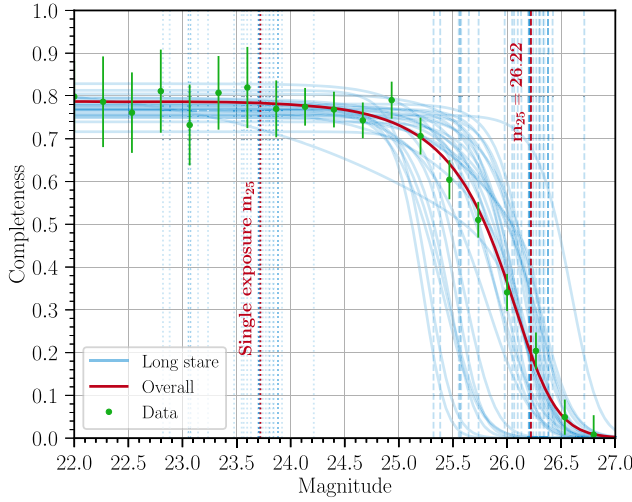
Figure 2. Left: completeness vs. magnitude for all long stares (solid blue) as well as the overall completeness for the survey (solid red), where all pointing groups are fit to the same probability curve. The vertical dashed lines indicate the best-fit m_{25} for each long stare (blue) and for the entire data set (red). The dotted lines represent the m_{25} each exposure in the long stare (blue) and the survey (red) would have if the data were not stacked. The green dots are the true binned recovery rates averaged over all long stares with objects limited to $r_{50,1} < r < r_{50,2}$ with Poisson error bars. Right: rate and magnitude for all synthetic sources (recovered in red and missed in blue) in our long stare with the deepest m_{25}. The shaded gray curves represent the completeness in Equation (7), and the horizontal dashed purple lines represent the 50% rate completeness terms ($r_{50,1}$ and $r_{50,2}$).

3.2. Joint Rate and Magnitude Selection Function

The selection function for each pointing group is not only a function of the source magnitude m, but also the rate of motion r, which we assume has the same functional form and parameters for all pointing groups. Thus, any fits of Equations (2) and (3) do not yield a meaningful completeness c, as in principle, one can add an arbitrary number of bright objects outside the rate range in which objects are meaningfully detected, or reduce the samples to a predefined range of rates. On the other hand, attempting to fit a *rate* completeness function requires the data to be limited to $m \ll m_{25}$, also necessarily reducing the sample size. We propose a methodology that accounts for both rate and magnitude selection functions simultaneously. We posit a simple smooth functional form for the rate completeness,

$$p(r|r_{50,1}, r_{50,2}, \kappa_1, \kappa_2, r_0)$$
$$= \begin{cases} \dfrac{1}{1 + \exp(\kappa_1(r - r_{50,1}))}, & r < r_0 \\[2ex] \dfrac{1}{1 + \exp(\kappa_2(r - r_{50,2}))}, & r > r_0, \end{cases} \quad (6)$$

corresponding to a logistic function in each end of the distribution (with $\kappa_1 < 0$ and $\kappa_2 > 0$), and a rate of 50% completeness for each side. The constant r_0 determines the transition between the two regimes, and can be chosen prior to the fit without loss of constraining power. In principle, we could attempt a similar statistical exercise as in the previous section to determine the optimal functional form for this function. In practice, the recovery efficiency drops to zero quite steeply, so we can safely posit this simpler form. We assume that this rate selection function is the same for all pointing groups, as the underlying digital tracking algorithm in `kbmod` does not change in each processing.

The combined rate and magnitude selection function becomes

$$p(m, r|m_{25}, k_1, k_2, r_{50,1}, r_{50,2}, r_0, \kappa_1, \kappa_2, c)$$
$$= c \cdot p(m|m_{25}, k_1, k_2, 1) \cdot p(r|r_{50,1}, r_{50,2}, \kappa_1, \kappa_2, r_0). \quad (7)$$

As before, c corresponds to the peak efficiency (when $m \ll m_{25}$ and $r_{50,1} < r < r_{50,2}$), the parameter range where objects are systematically recovered. To simplify the notation, we index each pointing group by μ and define $\boldsymbol{\theta}_\mu \equiv \{m_{25,\mu}, c_\mu, k_{1,\mu}k_{2,\mu}\}$, $\boldsymbol{\varphi} \equiv \{r_{50,1}, r_{50,2}\kappa_1, \kappa_2\}$ (noting that $\boldsymbol{\varphi}$ is independent of pointing group). We generalize the likelihood in Equation (1) by including the rate in addition to the magnitude of each implanted source and by multiplying it over all pointing groups, so

$$\mathcal{L} = \prod_\mu \mathcal{L}_\mu = \prod_\mu \left[\prod_\alpha p(m_\alpha, r_\alpha|\boldsymbol{\theta}_\mu, \boldsymbol{\varphi}) \prod_\beta [1 - p(m_\beta, r_\beta|\boldsymbol{\theta}_\mu, \boldsymbol{\varphi})] \right]. \quad (8)$$

Maximizing this $\mathcal{L}$ achieves our goal of obtaining a meaningful $\{c_\mu\}$ as well as characterizing the global rate selection function. One way of understanding Equation (8), then, is that for any given long stare, the rate selection function is informed by all other long stares, and so its constraints are stronger than what could be derived from any individual long stare, mitigating the effect of the joint dependency of magnitude selection function with the rates of the implanted sources. That is, we do not need as many sources in each (m, r) bin as we would if we were trying to reconstruct the joint rate and magnitude selection functions for each individual long stare. We illustrate this procedure in Figure 2.

4. Simulation Methodology

The characterization in the previous sections established that, for any given TNO inside our survey area, our discoverability

is a function of its on-sky motion (the rate) and magnitude. Thus, these are the primary parameters that our survey simulator needs in order to decide which objects are recovered. For any given object, the simulation proceeds as follows:

1. The object's orbital elements or state vector are integrated to all exposure times for the survey, and on-sky positions are generated. These positions are then used to check whether an object could be seen by a functional DECam CCD with this pointing. This properly accounts for chip gaps, and so this information is also used to decide whether this synthetic object would lie inside its search CCD for long enough to be properly recovered by the digital tracking methodology. It should be noted that only the synthetic objects that have successfully passed this check are included in the procedure of Section 3.

2. For each individual observations that could be recovered by the digital tracking procedure in a given long stare μ, we compute its rate r and mean magnitude m given the parent object's light curve (that can be constant), and sample a random uniform number $u \in [0, 1]$. If $p(m, r|\boldsymbol{\theta}_\mu, \phi) > u$, the object is considered to be recovered. Because $0 \leqslant p(m, r|\boldsymbol{\theta}_\mu, \phi) \leqslant c_\mu$, even bright objects have a $\sim 20\%$ chance of being discarded (as $\langle c_\mu \rangle \approx 0.8$), thus properly accounting for the area lost in the digital tracking procedure.

3. We compute the discoverability of the orbit given its arc (the time span between the first and last observations); the shortest arc after dropping either the first or the last night of data; and the number of nights in which the object was detected. These inform whether the object would be linked or not. For the processing of Smotherman et al. (2024), an object is linked with high efficiency (94%) if it was observed in at least four nights, with complete arcs spanning 0.8 yr (0.5 yr for the shortened arcs). We sample another uniform number $u \in [0, 1]$, and compare to the linking efficiency $\ell \approx 94\%$, to determine whether the object would be properly linked. This accounts for objects with the proper number of observations but whose detections are astrometric outliers (that is, the recovered positions do not agree with the nominal position for the object) or objects whose orbit fits yield poor results, as measured by their reduced χ^2 (see Smotherman et al. 2024 for a detailed discussion).

This simulation is implemented as an extension of DESTNOSIM (Bernardinelli et al. 2022), the survey simulator initially developed for the Dark Energy Survey (DES) search for TNOs (Bernardinelli et al. 2020a, 2022), and it makes use of the dynamics package Spacerocks (Napier 2020) to produce the ephemerides for the synthetic objects. The modular nature of the software allows characterization of any processing of DEEP, as well as any similar survey, such as DES. A full tutorial, as well as documentation and scientific applications of the software, are publicly available on a GitHub[20] repository. We encourage users to extend the capabilities of the software, as well as to publicly release their analyses in addition to their publications, ensuring that their results are reproducible by the community.

5. Applications

As an initial demonstration of the software, we study two simple scenarios with toy populations that quantitatively illustrate its simulation capabilities.
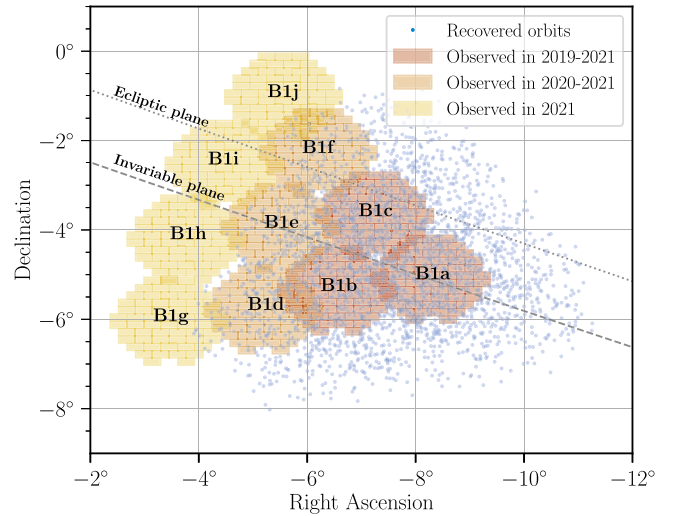
Figure 3. Sky coordinates of the B1 fields. The shaded regions indicate the observing coverage of each pointing, with the colors representing the observing years that covered each pointing (red: 2019–2021, orange: 2020–2021, yellow: 2021). The blue dots show a representative sample (20%) of the TNO that could be successfully linked and are displayed in their nominal sky position in 2020 January 1 (the reference epoch). The dotted (dashed) line mark the location of the ecliptic (invariable) plane. We note that few objects in fields B1g–B1j are recovered: this is expected, as these fields were not observed in 2019 and 2020, so objects from these fields have not yet been imaged to be linkable in the data.

5.1. Effective Search Area

The 2019–2021 data from the DEEP B1 field include 10 distinct DECam pointings (see Figure 3). Using the camera's nominal field of view (3 deg²), it would seem that our search area for multiyear orbits is $\approx 30 \, \text{deg}^2$. This is, however, not precise: not all fields were observed in all years, not all of DECam's CCDs are functional, and there is a fraction of area lost due to the survey's observing geometry (the combination of fields, cadence, and camera footprint). The observing strategy of DEEP is such that fields are imaged in an expanding "cone" pattern: in 2019, the B1a, B1b, and B1c fields were visited; in 2020, the B1a through B1f fields were observed, and all fields were observed in 2021.

We use the simulator to estimate the "effective" search area, that is, the area in the sky where any potential bound orbit could be observed by DEEP and then subsequently be recovered by the linking procedure. For these purposes, we ignore the magnitude and rate selection functions, as our interest is in the geometrical design of the survey. We simulate an isotropic, all-sky population of objects as defined above, but we fix their initial distance to 40 au (it should be noted that, as defined here, this effective area is distance dependent, as we are accounting for the motion of the object). The procedure of Bernardinelli et al. (2020a) distributes orbits in a spherical Fibonacci lattice, and so the fraction of the points we recover is a proxy for the area on the sphere that is available for objects to be searched (see González 2010 for a detailed explanation of these lattices). We simulated 4×10^7 TNOs, and of these, 14,408 satisfy the linking criteria of Section 4, leading to $A_{\text{eff}} \approx 14.8 \, \text{deg}^2$. We emphasize that this effective area will change as the survey progresses.

5.2. Limits on a Distant, Dynamically Cold Population

There are few objects known in the outer Kuiper Belt, defined as having a larger than Neptune's 2:1 mean motion

---

resonance and low eccentricities ($e \lesssim 0.24$, Gladman & Volk 2021). While the high-inclination component of this population extends as far as the 3:1 resonance (Sheppard et al. 2016; Bernardinelli et al. 2022), this component is thought to be created by a combination of mean motion resonances and the Kozai effect (Gomes et al. 2008), and no dynamically cold ($e \leqslant 0.24$ and $i \leqslant 5°$) objects with $a > 60$ au are known. In Smotherman et al. (2024), we recovered 110 TNOs, primarily cold Classicals (low-inclination, low-eccentricity objects between Neptune's 3:2 and 2:1 mean motion resonances; see Figure 7 of Smotherman et al. 2024). In agreement with other surveys, we also also did not detect any objects in this high-a, low-e, and low-i region of the parameter space, so we can use our simulator to place limits on the existence of a distant thin disk of objects past this a range. We emphasize that this is not the first time such limits have been placed (Trujillo & Brown 2001; Trujillo et al. 2001; Allen et al. 2002; Bernstein et al. 2004), but no limits with our combination of area and magnitudes have been placed to date, and such a population has not been entirely ruled out (see Section 6.2.3 of Gladman & Volk 2021).

We simulate a population similar to the stirred classical Kuiper Belt population of Petit et al. (2011), following their $a \propto a^{-5/2}$ surface density for the Kuiper Belt, but moving it further to the $60 \leqslant a \leqslant 80$ au range. The other parameters follow Petit et al. (2011): e uniformly distributed between 0 and 0.03; $i \propto \sin(i)\exp(-i^2/(2\sigma_i^2))$, where $\sigma_i = 2°.6$ and ω, Ω, $\mathcal{M}$ uniform in the $0°$–$360°$ range. We simulate this population in the $24 \leqslant m \leqslant 26$ range, corresponding roughly to $5 < H < 8$ at these distances. It is worth noting that, for $m \ll 24$, our magnitude efficiency is constant, and our population estimates are limited by our survey area, so we do not need to simulate brighter objects. We measure the fraction of members of this population recovered at each magnitude bin m_b. This allows us to compute our probability of detecting an object $p(m_b)$. Assuming a size distribution, we can convolve it with our detection probability $p(m_b)$ to place limits on the number of objects with an absolute magnitude $H \leqslant H_b$. Using Poisson statistics, we have that an undetected population whose simulation leads to three detections on average can be rejected it with 95% confidence, as the Poisson error bar is $\sqrt{3}$ for such a process. Assuming the Fraser et al. (2014) size distribution for the cold Classical population, namely, a broken power law with slope 1.5 for the larger objects and 0.3 for the smaller objects, with the break at $H = 6.9$, we have $N(H \leqslant 8) = 8 \times 10^3$ objects in this region of the outer Solar System. It is important to note that any size distribution that is *shallower* than that of Fraser et al. (2014) would lead to a stronger constraint, that is, a smaller number of objects. A more useful limit would be for $a \sim 90$ au (Gladman & Volk 2021), but our completeness drops sharply after $d \approx 80$ au (see Figure 1). Further data releases of the DEEP survey will attempt to recover and place limits to these more distant populations.

6. Summary

We constructed a survey simulator for the DEEP survey via injection of a population of synthetic outer Solar System objects that exhausts the parameter space of orbits at Kuiper Belt distances (and beyond) into the DEEP images. This software, publicly released on `GitHub`, enables the user to characterize our observational biases as a function of magnitude and rate of motion for each pointing group, and of the survey as a whole. Such a construction enables further studies of recoverability of populations of trans-Neptunian objects by the survey, and thus detailed comparisons between models of the structure of the trans-Neptunian region and the TNOs observed by the DEEP survey. In addition to the software release, we will provide real-world use cases in the form of Jupyter Notebooks, ensuring that our analyses are reproducible by the community.

In this first release, we focused on the 2019–2021 data from the B1 fields (Smotherman et al. 2024) processed through `kbmod` (Whidden et al. 2019; Smotherman et al. 2021), but the methodology developed here is general and can be extended to any survey that uses digital tracking to recover TNOs. Our lack of dynamically cold objects in distant ($a > 60$ au) orbits, combined with the survey simulator, enables us to derive that this yet-undetected potential population has at most 8×10^3 members with $H \leqslant 8$.

ORCID iDs

Pedro H. Bernardinelli ● https://orcid.org/0000-0003-0743-9422
Hayden Smotherman ● https://orcid.org/0000-0002-7895-4344
Zachary Langford ● https://orcid.org/0000-0001-7574-4440
Stephen K. N. Portillo ● https://orcid.org/0000-0001-8132-8056
Andrew J. Connolly ● https://orcid.org/0000-0001-5576-8189
J. Bryce Kalmbach ● https://orcid.org/0000-0002-6825-5283
Steven Stetzler ● https://orcid.org/0000-0002-7712-6678
Mario Jurić ● https://orcid.org/0000-0003-1996-9252
William J. Oldroyd ● https://orcid.org/0000-0001-5750-4953
Hsing Wen Lin (林省文) ● https://orcid.org/0000-0001-7737-6784
Fred C. Adams ● https://orcid.org/0000-0002-8167-1767
Colin Orion Chandler ● https://orcid.org/0000-0001-7335-1715
Cesar Fuentes ● https://orcid.org/0000-0002-5211-0020
David W. Gerdes ● https://orcid.org/0000-0001-6942-2736
Matthew J. Holman ● https://orcid.org/0000-0002-1139-4880
Larissa Markwardt ● https://orcid.org/0000-0002-2486-1118
Michael Mommert ● https://orcid.org/0000-0002-7817-3388
Kevin J. Napier ● https://orcid.org/0000-0003-4827-5049
Matthew J. Payne ● https://orcid.org/0000-0001-5133-6303
Darin Ragozzine ● https://orcid.org/0000-0003-1080-9770
Andrew S. Rivkin ● https://orcid.org/0000-0002-9939-9976
Hilke Schlichting ● https://orcid.org/0000-0002-0298-8089
Scott S. Sheppard ● https://orcid.org/0000-0003-3145-8682
Ryder Strauss ● https://orcid.org/0000-0001-6350-807X
David E. Trilling ● https://orcid.org/0000-0003-4580-3790
Chadwick A. Trujillo ● https://orcid.org/0000-0001-9859-0894

References

Allen, R. L., Bernstein, G. M., & Malhotra, R. 2002, AJ, 124, 2949
Bannister, M. T., Gladman, B. J., Kavelaars, J. J., et al. 2018, ApJS, 236, 18
Bannister, M. T., Kavelaars, J. J., Petit, J. M., et al. 2016, AJ, 152, 70
Batygin, K., Adams, F. C., Brown, M. E., & Becker, J. C. 2019, PhR, 805, 1
Belyakov, M., Bernardinelli, P. H., & Brown, M. E. 2022, AJ, 163, 216
Bernardinelli, P. H., Bernstein, G. M., Montet, B. T., et al. 2021, ApJL, 921, L37
Bernardinelli, P. H., Bernstein, G. M., Sako, M., et al. 2020a, ApJS, 247, 32
Bernardinelli, P. H., Bernstein, G. M., Sako, M., et al. 2020b, PSJ, 1, 28
Bernardinelli, P. H., Bernstein, G. M., Sako, M., et al. 2022, ApJS, 258, 41
Bernstein, G. M., Trilling, D. E., Allen, R. L., et al. 2004, AJ, 128, 1364
Flaugher, B., Diehl, H. T., Honscheid, K., et al. 2015, AJ, 150, 150
Fraser, W. C., Brown, M. E., Morbidelli, A., Parker, A., & Batygin, K. 2014, ApJ, 782, 100
Gladman, B., Kavelaars, J. J., Nicholson, P. D., Loredo, T. J., & Burns, J. A. 1998, AJ, 116, 2042
Gladman, B., & Volk, K. 2021, ARA&A, 59, 203
Gomes, R. S., Fernández, J. A., Gallardo, T., & Brunini, A. 2008, in The Solar System Beyond Neptune, ed. M. A. Barucci et al. (Tucson, AZ: Univ. Arizona Press), 259
González, A. 2010, MaGeo, 42, 49
Jewitt, D., & Luu, J. 1993, Natur, 362, 730
Jones, R., Gladman, B., Petit, J. M., et al. 2006, Icar, 185, 508
Jurić, M., Kantor, J., Lim, K. T., et al. 2017, in ASP Conf. Ser. 512, Astronomical Data Analysis Software and Systems XXV, ed. N. P. F. Lorente, K. Shortridge, & R. Wayth (San Francisco, CA: ASP), 279
Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111
Kass, R. E., & Raftery, A. E. 1995, JASA, 90, 773
Kavelaars, J. J., Petit, J. M., Gladman, B., et al. 2021, ApJL, 920, L28
Lawler, S. M., Kavelaars, J. J., Alexandersen, M., et al. 2018a, FrASS, 5, 1
Lawler, S. M., Shankman, C., Kavelaars, J. J., et al. 2018b, AJ, 155, 197
Morbidelli, A., & Nesvorný, D. 2020, The Trans-Neptunian Solar System (Amsterdam: Elsevier), 25
Napier, K., 2020 SpaceRocks, https://github.com/kjnapier/spacerocks
Napier, K. J., Gerdes, D. W., Lin, H. W., et al. 2021, PSJ, 2, 59
Napier, K. J., Lin, H. W., Gerdes, D. W., et al. 2024, PSJ, 5, 50
Petit, J. M., Kavelaars, J. J., Gladman, B. J., et al. 2011, AJ, 142, 131
Schwarz, G. 1978, AnSta, 6, 461
Shankman, C., Kavelaars, J. J., Bannister, M. T., et al. 2017, AJ, 154, 50
Sheppard, S. S., Trujillo, C., & Tholen, D. J. 2016, ApJL, 825, L13
Smotherman, H., Connolly, A. J., Kalmbach, J. B., et al. 2021, AJ, 162, 245
Smotherman, H., Bernardinelli, P. H., Portillo, S. K. N., et al. 2024, AJ, 167, 136
Strauss, R., Trilling, D. E., Bernardinelli, P. H., et al. 2024, AJ, 167, 135
Towns, J., Cockerill, T., Dahan, M., et al. 2014, CSE, 16, 62
Trilling, D. E., Gerdes, D. W., Juric, M., et al. 2024, AJ, 167, 132
Trujillo, C., Fuentes, C., Gerdes, D. W., et al. 2024, AJ, 167, 133
Trujillo, C. A., & Brown, M. E. 2001, ApJL, 554, L95
Trujillo, C. A., Jewitt, D. C., & Luu, J. X. 2001, AJ, 122, 457
Whidden, P. J., Bryce Kalmbach, J., Connolly, A. J., et al. 2019, AJ, 157, 119