UPDATED DIGEST2 – THE NEO CLASSIFICATION CODE Peter Vereš[1], R. Cloete[1], R. Weryk[2], A. Loeb[1], M. J. Payne[1], [1]Harvard-Smithsonian Center for Astrophysics, Minor Planet Center. 60 Garden Street, Cambridge, MA 02138, USA; pveres@cfa.harvard.edu; [2]Physics and Astronomy, The University of Western Ontario,1151 Richmond Street, London ON N6A 3K7, Canada

Keywords: *NEO identification,short-arc orbit classification, Minor Planet Center, digest2*

Introduction: Traditionally, Near-Earth Objects (NEOs) have been discovered due to their rapid motion and the immediate follow-up analysis that enables quick orbit determination and second-night recovery. However, with the advent of wide-field asteroid surveys and the need to optimize valuable follow-up time, it is crucial to differentiate between the background population of main-belt objects and those that are characteristic of NEOs. For almost two decades, the digest2 classification code [1] has been utilized by the Minor Planet Center and the NEO community to score short-arc unidentified tracklets, which are then posted to the Near-Earth Object Confirmation Page for immediate follow-up. The input for digest2 is the MPC's 80-character astrometry format, which describes the astrometric positions of NEO candidates. The software works as a binary NEO classifier, but there are known disadvantages that lead to occasional cases where digest2 fails to flag NEOs as interesting [2]. Among the most prominent problems are (i) the synthesis of a two-detection tracklet from the submitted observations, which results in both the loss of information embedded in the apparent motion if more than two positions are provided, and (ii) the astrometric errors that are assumed for each observatory code.

Table 1: Fraction of ADES-submitted astrometry as a function of observation time and the total number of astrometric positions published by MPC or present in the isolated tracklet file.

Year	Fraction of ADES	Observations
2018	13%	23.9 million
2019	40%	33 million
2020	48%	42.6 million
2021	82%	32 million
2022	91.5%	34 million

In this work, we present an update to the digest2 code that: (i) ingests astrometry from the new astrometric format - ADES [3] - allowing the submitter to provide astrometric uncertainties for each measured astrometric position (Table 1), and (ii) implements a method for curvature computation of tracklets with at least three detections and astrometric uncertainties (N.B. statistically significant deviation from the great circle fit of a short-arc tracklet suggests the object is very close to Earth regardless of its rate of motion, and thus its digest2 NEO score should be increased). We have also updated the population model used in the digest2 code, and updated the assumed per-observatory error model.

Data sources:

For our analysis, we selected four independent data sources (Table 2): the first two were anonymized tracklets selected from known NEOs and Main-belt asteroids (MBAs). Using known orbits serves as a true-positive and false-positive indicator. We selected tracklets containing at least three detections with a magnitude threshold of fainter than $V = 19.5$. The magnitude threshold was selected because the main-belt population is roughly complete to the given apparent magnitude, and all brigher objects are likely either NEOs, comets, distant objects or interesing enough to be flagged based only on their magnitude. We also used the abslute magnitude threshold (H): $H > 20$ for NEOs and $H > 13$ for non-NEOs.

In addition, we explored a set of isolated tracklet file (ITF) tracklets, also with at least three detections and without a magnitude limit.

The last data set that we studied contained NEOCP discovery tracklets, regardless of their magnitudes or final attribution. We found that about 55% of NEOCP tracklets are actually NEOs, 30% non-NEOs, and the remaining tracklets were either undefined, artifical, or deleted by the observer.

Table 2: Number of tracklets, detections and time-range of data samples. Only about 70% of NEOCP data had ADES information.

Type	Tracklets	Observations	Date-range
NEOs	30,829	110,989	Jan 2020-Jan 2023
MBAs	873,239	3,430,460	Jan 2020-Jan 2023
ITF	97,693	373,202	Jan 2020-Jan 2023
NEOCP	22,100*	83,125*	Feb 2019-Feb 2023

Software updates:

The digest2 software requires a population model in the form of binned data in four dimensions: three orbital elements (perihelion distance q, eccentricity

e, inclination i) and absoute magnitude (H). There are two populations models. The first is the full population model, which represents the complete Solar System to a given H, and is based on the synthetic Solar System model by [4] which is comprised of over 14 million Keplerian orbital elements. This population model is used for computing so-caller 'Raw' digest2 scores. The second model represents the undiscovered portion of the Solar System, derived by subtracting the binned full Solar System Model from the catalog of discovered objects, e.g. from MPCORB[1]. Undiscovered populations yield a 'NOID' digest2 score. We updated the population mode based on the MPCORB catalog from Feb 26, 2023.

Digest2 also relies on the astrometric uncertainties that are used to "dither" the end-points of the synthesized tracklets (see [1]). Traditionally, the end points were dithered based on a representative uncertainty value for a given observatory code. We extended the list of observatory codes with assigned uncertainties to 143 based on the orbit fits to all known objects from MPCORB by orbfit[2].

A major update of the digest2 code is the added support for reading the astrometry in the ADES format, submitted in eXtensible Markup Language (XML). That is, the previous method of providing digest2 input was through the 80-column MPC1992 format. However, the new format allows for the submission of substantially more information for each tracklet, including measured astrometric uncertainties for each tracklet instead of assumed values. To make use of this new format, digest2 ADES-formatted files must have the ".xml" suffix. If the input file does not have this suffix, the code will assume it is an MPC1992 input file.

Another important update is that the digest2 code now includes the computation of curvature within three or more detection tracklets. When an object is in close proximity to Earth, its motion may deviate from a simple great-circle path, even over a short period of time. To account for this, we have added a new parameter, denoted by RMS', which is calculated from the astrometric uncertainties provided in the XML file, or from the per-obscode uncertanties specified in the configuration file.

Results:

We compared the performance of the updated digest2 code, which includes a new population model and improved astrometric uncertainties, to the previous version of the code. Additionally, we evaluated the performance of the updated digest2 using the new ADES/XML input, in comparison to the traditional MPC1992 input. Our assessment was based on the number of true-positive and false-positive Near-Earth Object (NEO) identifications, using Raw and NOID digest2 scores at the current threshold of $D = 65$.

Our results show that in both cases, the updated digest2 code provide a better true-positive and false-negative NEO identification. Also, the XML-enhanced digest2 code is significantly faster than the previous version reading the MPC1992 format (Table 3).

Table 3: Runtime comparisons between MPC1992 and ADES XML formats, in minutes, on the data sets from Table 1

Type	MPC1992	ADES XML	Improvement factor
NEOs	3:20	0:37	5.4
MBAs	251:18	39:29	6.3
ITF	26:43	4:12	6.3
NEOCP	1:58	0:20	5

We also explored the possibility of detecting in-tracklet curvature by RMS' and finding NEOs among the low-scoring digest2 tracklets. Our findings suggest curvature exists for a few known NEOs but also yielded more false-positives. Closer examination suggests that the reported uncertainties were likely underesimated.

In summary, (i) the updated population model ensures that the digest2 score will weight the tracklets accurately based on the current undiscovered population. We emphasize the importance of regular population model updates, particularly when the number of discovered objects increases rapidly. (ii) Our updated digest2 software is faster, and allows for the detection of tracklet curvature. We encourge observers to submit their astrometric data in the XML ADES format and provide accurate astrometric and timing uncertainties for each detection.

References: [1] S. Keys, et al. (2019) 131(1000):064501 doi.arXiv:1904.09188. [2] R. Wainscoat, et al. (2022) 373:114735 doi. [3] S. R. Chesley, et al. (2017) in *AAS/Division for Planetary Sciences Meeting Abstracts #49* vol. 49 of *AAS/Division for Planetary Sciences Meeting Abstracts* 112.14. [4] T. Grav, et al. (2011) 123(902):423 doi.

[1] https://www.minorplanetcenter.net/iau/MPCORB.html

[2] http://adams.dm.unipi.it/~orbmaint/orbfit/