# Statistical analysis of a weighting scheme for asteroid observation astrometric errors taking into consideration the classification of the observed asteroids

**Nicolò Stronati**[a*], **Laura Faggioli**[b], **Marco Micheli**[b], **Marta Ceccaroni**[a]

[a] *Cranfield University, Cranfield, England MK43 0AL, United Kingdom*
[b] *ESA NEO Coordination Centre, Largo Galileo Galilei, 1, 00044 Frascati, Italy*
[*] Corresponding Author, nicolo.stronati@cranfield.ac.uk

**Abstract**

Observations of asteroids and other near-Earth objects are of great importance for planetary defence activities, the purpose of which is to determine their positions in space and the probabilities of Earth impacts, as well as developing strategies to mitigate this risk. In this framework, having precise observations is important to describe accurately the orbits of near-Earth asteroids. However, given a general absence of a-priori uncertainty information, the single observations are given proper weights that reflect the accuracy expected by the observers who perform the observations. The weights are calculated for each observer on the base of statistical analysis of systematic and random errors and providing them with an accurate definition is necessary if the magnitude of the error of a single observation is to be correctly estimated. In this paper a statistical analysis on the residuals of the astrometric data provided by the major surveys is presented introducing a dynamic classification of observed asteroids. The observations are thus subdivided between those relative to Near Earth and Main Belt Asteroids and the quality of the data for each station is studied focussing on this classification. The results show that most of the considered stations have the same quality regardless of the measured object, while four of them show a dependency on this factor.

## 1. Introduction

The Asteroids Orbit Determination (OD) process requires astrometric observations of the given object that are then integrated through numerical algorithms based on the classic least-square method. An observation provides information on the position in space of a celestial body at a given epoch and the integration of the all available astrometric data, applying the proper dynamic model, define our best possible guess of the object's real trajectory. The first attempt for OD relies on just three observations while, with the addition of more observations, the orbit can be refined by finding a least-square solution of a perturbed N-body problem.

Observations are subject to errors, both systematic and random, that lead to uncertainties in the OD process and its outcome will be included in an uncertainty region defined by the overall uncertainties. Many observation biases can affect the results and move the error distribution away from the expected Gaussian distribution, in agreement with the Central Limit Theorem [1].

Astrometric errors management has become a branch of study that runs in parallel with the technological improvement of observation tools. For it, the analysis of a high amount of data, i.e. observations residuals, is necessary, as per the usual probabilistic approach to the matter. The main errors that can affect the measure are generally due to biases in the star catalogues, used for the determination of the position of the object at the moment of the observation [2], and the observing stations; strategies to correct both these sources have been analysed and implemented.

Some studies have presented methods to debias the catalogue error [3, 4] reducing all the data obtained with older and less precise catalogues with respect to the recent 2MASS and Gaia DR2 providing an enhancement of the previous observations.

The stations are however responsible for the quality of the instruments, which include telescopes, focal lengths and pixel resolution, the seeing of the station, characteristic of the location, that limits the angular resolution, the number of observations performed during one night, the precision in time and position. These factors can affect the quality of the astrometric data. To account all these sources of uncertainty a statistical treatment based on an ad-hoc weighting scheme is performed on the astrometric data that reflects the quality of the observation.

The usual OD pipeline adopts an observations weighting scheme based on the overall performance of the observatory [2, 3], which takes as a parameter the Root Mean

Square (RMS) of the O-C (Observed-Computed) residuals. The idea under this assumption is that the equipment, the technical precision, the methodology change station by station and must be considered when assessing the weights.

Moreover, it has also been suggested that, within the same station, the astrometric performance are likely to vary also due to other factors, such as the epoch when the observation is done, or physical parameters such as the light curve of the observed object [5]. In fact, it is expected that over the time small or massive improvements can be done to instruments, software and also the introduction of the already mentioned less biased catalogues for the data reduction. Furthermore, other aspects have been analysed in previous studies, including, among other, the magnitude of the observed objects and their motion rate.

The aim of this work is to enhance the observations weighting system, by analysing the possible biases introduced on astrometric residuals from the dynamic features of the observed object, by dividing them between Near Earth Asteroids (NEAs), Main Belt Asteroids (MBAs), Trans-Neptunian Objects (TNOs). This is justified by considering the big differences shown by these objects when observed from planet Earth in, among others, motion rates and number of available oppositions. Said that, clustering them and providing statistical results for each of the classes can help develop a new weighting system able to enhance the quality of the available data and the OD process.

Following the baseline presented in a previous work [5], we will analyse the performances of the most productive surveys, active or no more in operation, differentiating between the already described asteroid classes.

The structure of this paper can be summarized as follows: in Section 2 the methodology employed to select and gather the data is described; Section 3 analyses the results of the RMS computation for the considered stations.

## 2. Data collection

As of February 2022 the Minor Planet Center orbit (MPCORB) database* accounted $1177443$ asteroids, of which $612011$ numbered. The observations of these bodies can be gathered from the ESA's Near-Earth Objects Coordination Centre (NEOCC)† database, where, for most of the objects on the MPCORB index, one can access the corresponding .rwo file. In these files each line corresponds to a single observation, i.e. the observation epoch and the angular coordinates given in terms of declination (DEC) and right ascension (RA), that provide the position of the object with respect to the station. This is uniquely identified by its MPC code‡, from which it is automatic to trace the location of the station that is performing the measure. In the following sections we will often refer to the stations with their MPC code for the sake of brevity.

The observation lines in the .rwo file also contain other information on the measure such as the accuracy and the (O-C) residuals, indicated for both the DEC and RA. The *computed* trajectory is calculated by the NEOCC [6] with the least square method and the proper dynamical model applied, taking into account the observations filtered after an outlier rejection process.

In this paper the subdivision of the observations to be analysed for the RMS computation among three different asteroids macro-families identified by the distance of their orbit by the Sun is of central importance. The size of the perihelion of each asteroid has been considered to discriminate whether an asteroid is part of NEO, MBA or TNO class. Although there are not objective and fixed boundaries for the definition of these families, the asteroids in the MPCORB database have been allocated in a specific class following these conditions on the perihelion distance $q$:

- the asteroid is a NEO if $q < 1.3$ au;
- the asteroid is a MBA if the value of $q$ is between $1.78$ and $5$ au;
- the asteroid is a TNO if $q > 30$ au.

Logically, all the asteroids not included in this scheme will not be taken in consideration for the following analyses as they are not part of the families of interest.

A criterion on the number of oppositions for which the asteroids have been observed has been applied too. In order to consider objects with a well determined orbit only asteroids with strictly more than two oppositions have been included. This is because the more the oppositions, the more the available observations are spaced in time, the more the OD pipeline gives reliable results, making the RMS statistics on the residuals more significant. Furthermore, this will allow us to include both numbered and many non numbered asteroids, in order to have the highest amount of observations possible.

### *2.1 RMS computation*

For each of the stations listed in the first column of Table 1, the (O-C) residuals are collected as of May 2022 and the RMS of residuals are computed. The results by station, by class of asteroid, in RA and DEC are shown together

*https://minorplanetcenter.net/iau/MPCORB.html

†https://neo.ssa.esa.int/

‡https://minorplanetcenter.net/iau/lists/ObsCodesF.html

Table 1: RMS of the residuals in RA and DEC for 17 of the most productive surveys subdivided among the different asteroids macro-families. The number of available obervations for each sation for each class of asteroid is shown.

| Obs code | NEO | | | MBA | | | TNO | | |
|---|---|---|---|---|---|---|---|---|---|
| | $RMS_{RA}$ ($''$) | $RMS_{DEC}$ ($''$) | N. obs | $RMS_{RA}$ ($''$) | $RMS_{DEC}$ ($''$) | N. obs | $RMS_{RA}$ ($''$) | $RMS_{DEC}$ ($''$) | N. obs |
| 704 | 0.619 | 0.610 | 95688 | 0.650 | 0.651 | 30963271 | 0.220 | 0.310 | 8 |
| G96 | 0.287 | 0.266 | 110097 | 0.288 | 0.265 | 53423691 | 0.266 | 0.319 | 71 |
| F51 | 0.119 | 0.114 | 99469 | 0.104 | 0.102 | 52320580 | 0.130 | 0.128 | 80067 |
| F52 | 0.106 | 0.104 | 26898 | 0.098 | 0.100 | 13334711 | 0.113 | 0.107 | 4345 |
| 703 | 0.592 | 0.553 | 128695 | 0.628 | 0.591 | 30902815 | 0.824 | 0.610 | 11 |
| 691 | 0.437 | 0.394 | 28539 | 0.346 | 0.276 | 13267897 | 0.551 | 0.543 | 315 |
| G45 | 0.332 | 0.330 | 28980 | 0.335 | 0.340 | 13463362 | N/A | N/A | 0 |
| 699 | 0.596 | 0.547 | 10351 | 0.614 | 0.564 | 5080913 | N/A | N/A | 0 |
| 644 | 0.337 | 0.398 | 9603 | 0.281 | 0.345 | 3781872 | 0.345 | 0.370 | 796 |
| D29 | 0.398 | 0.388 | 15239 | 0.419 | 0.408 | 9717363 | 0.465 | 0.441 | 1156 |
| C51 | 0.529 | 0.570 | 37586 | 0.604 | 0.682 | 3303828 | N/A | N/A | 0 |
| E12 | 0.375 | 0.387 | 14078 | 0.468 | 0.496 | 2084641 | N/A | N/A | 0 |
| 608 | 0.592 | 0.650 | 5334 | 0.585 | 0.716 | 1147870 | 0.379 | 0.391 | 85 |
| J75 | 0.402 | 0.389 | 2030 | 0.423 | 0.394 | 1111033 | N/A | N/A | 0 |
| I41 | 0.312 | 0.273 | 31544 | 0.212 | 0.152 | 13185151 | 0.126 | 0.113 | 666 |
| T05 | 0.316 | 0.312 | 73523 | 0.339 | 0.337 | 19077215 | 0.275 | 0.265 | 1947 |
| T08 | 0.323 | 0.336 | 78868 | 0.347 | 0.363 | 20375204 | 0.300 | 0.329 | 2553 |

with the total number of observations $N$ collected and used for the computation of the RMSs.

The number of collected data shows the statistical degree of confidence of the results and give information on how the available amount of observations are distributed among the different classes of objects. It stands out that MBAs can rely on a much bigger sample of data than for the other two macro-families, due to the facts that they are the highest minor planets in number, almost the 92% of the total[§], and that given their synodic period with the Earth, they can benefit of more oppositions.

On the other hand, just few observations are available for the TNO class: there are even five stations (G45, 699, C51, E12, J75), all equipped with small-aperture telescopes, that have no data at all on objects further than 30 au from the Sun and the most productive survey for TNO is F51 with slightly more than 80000 observations. So, given far less data available for this family a comparison with the results of MBA and NEO will not be performed as it would not be statistically meaningful.

[§] https://minorplanetcenter.net/mpc/summary

## 3. Statistical analysis for the different asteroid classes

From the results in Table 1 the 17 considered stations present different values of RMS of residuals and hence different measure accuracy and precision. In Figure 1 these results are illustrated for the values in RA and DEC and one can easily see how the global accuracy changes not only with respect to the station, but also depending on the class of asteroid.

However, for most stations, the differences between the RMS of residuals for NEAs and MBAs are enough small to be considered within the signal noise, even if, for most of them, the best accuracy seems to be achieved by the observation of NEAs.

There are two stations for which the RMS values differ, between the two classes of asteroids, for less than $0.08''$: G96 (Mt. Lemmon Survey) and F52 (Pan-STARRS 2, Haleakala). On the contrary, four stations in particular present the highest gap when changing the type of observed object.

E12 (Siding Spring Survey) and C51 (WISE) present a higher accuracy for NEAs, while I41 (Palomar Mountain) for MBAs: these results can be easily explained by the scopes and working procedures of the three survey. E12,

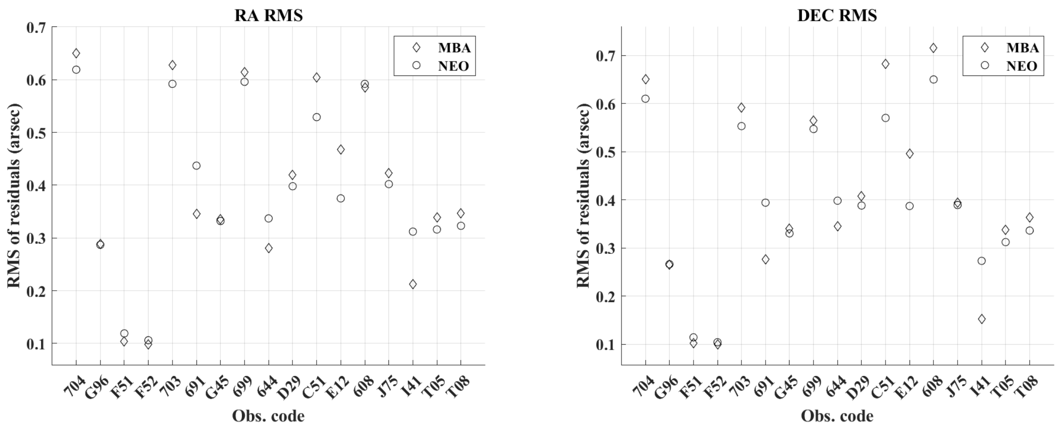


Fig. 1: RMS of resdiuals in RA (left panel) and DEC (right panel) for all the stations, considering the values for NEAs and MBAs

currently no longer operational, was known to manually remeasure NEAs detections which justifies the better accuracy. Similarly WISE (Wide-field Infrared Survey Explorer), has from 2013 the target of providing, under the NASA project Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), observations for NEAs. For I41, as it is not purely dedicated to Near Earth Objects, or asteroids in general [7], it might be expected that, when a relatively faster moving object like a NEA is observed the result of the measure can lead to a slightly worse accuracy.

On the contrary, 691 (Steward Observatory, Kitt Peak-Spacewatch), which is still providing data since 1984, provides better observations for MBAs, and a deeper statistical analysis is presented below to investigate the accuracy and precision in its measurements.

Finally, the surveys that guarantee the best accuracy are F51 and F52, the two Pan-STARRS stations on Haleakala. The observations quality is also similar as they share the same location, instruments, equipment, searching algorithms, even if F52 presents slightly better performances. In fact, it has been developed and installed subsequently to F51, taking advantage of its findings, fact that allowed a general improvement of the components for the most recent telescope [8].

These results confirm those presented in the baseline work [5] and add information on F52, for which enough data were not available at the time when the aforementioned work was presented.

### *3.1 Statistical analysis of the astrometric data from 691 (Steward Observatory, Kitt Peak-Spacewatch)*

From previous results it turned out that station 691 (Steward Observatory, Kitt Peak-Spacewatch), although it is a long in service and highly productive survey, shows an important difference in the RMS of residuals between NEAs and MBAs. In this section a statistical analysis of the residuals for this observatory is presented, to understand how this difference can be interpreted.

In Figure 2 the histograms of the residuals are shown for the observations in RA and DEC, always paying attention to compare the values for MBAs and NEAs, superimposing the two diagrams.The curves representing the relative Normal Gaussian Distributions have also been plotted, based on the mean and standard deviation values computed for the data and listed in Table 2. Both the histogram and the gaussian curve have been normalized according to the statistical definition of the probability density function (pdf).

Table 2: Avarage ($\overline{res}$) and standard deviation ($\sigma$) of the astrometric data available for station 691 in RA and DEC, divided between the two classes of asteroids MBA and NEA.

| Ast. Class | $\overline{res}_{RA}$ ($''$) | $\sigma_{RA}$ ($''$) | $\overline{res}_{DEC}$ ($''$) | $\sigma_{DEC}$ ($''$) |
|---|---|---|---|---|
| MBA | $-0.1454$ | $0.3137$ | $-0.0260$ | $0.2747$ |
| NEA | $-0.0847$ | $0.4284$ | $-0.0691$ | $0.3883$ |

If the data were distributed following an ideal Normal

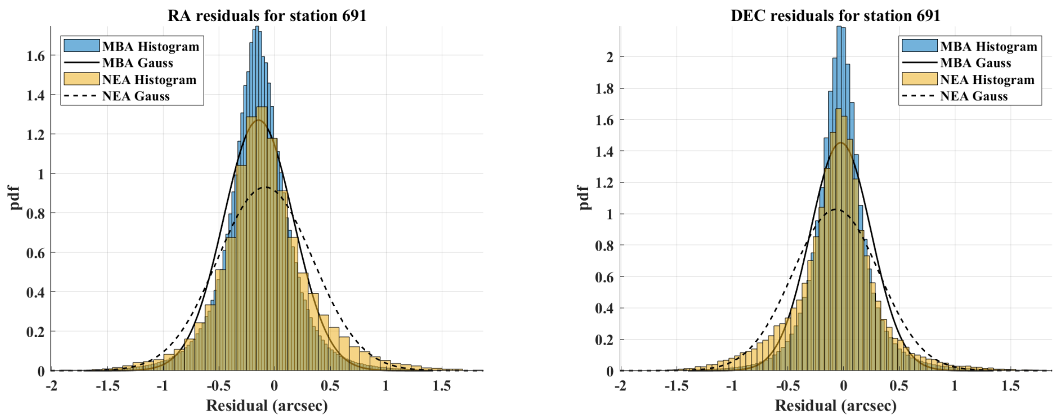


Fig. 2: Probability density function (pdf) of the distribution of the residuals for station 691 in Right Ascension (left panel) and Declination (right panel). The histogram and the Gaussian Normal Distribution are represented superimposed for the residuals of MBA and NEA observations distinctly.

Distribution, with just random observation errors, the expected average of the residuals should be zero for all the cases. Actually, many external factors can contribute to generate biases or systematic errors, which lead to non-zero averages. This non-zero average, that can be seen as the global bias of the station, is overall smaller for the measures in DEC and raises for those in RA. The standard deviation values for the four cases are almost equal to the RMS values listed in Table 1, which is a typical characteristic of a data set with zero, or almost zero, as in this occasion, average.

A parameter that can help understand the bias of the station towards the measures for the two classes of asteroid is the $(\overline{res}/RMS)$ ratio, which for NEA is $-0.1938$ in RA and $-0.1754$ in DEC while for MBA $-0.4202$ in RA and $-0.2504$ in DEC. These values, together with those related to the RMSs and $\sigma$, may lead to the conclusion that this station is able to provide a better accuracy for observations of NEAs, but a higher precision, with a smaller standard deviation, for Main Belters. The difference, again, may be due to the effort put in the measure of the angular positions of objects belonging to the different classes or to objective difficulties of the instruments in providing the same precision for Near Earth Objects.

## 4. Conclusions

A statistical analysis of the astrometric residuals for the most productive surveys has been performed taking into account the asteroid dynamic macro-classes. The RMS of the residuals for both the position measures in Right Ascension (RA) and Declination (DEC) have been computed and compared, and turned out that most of the stations do not show differences in observation accuracy when observing objects belonging to different classes.

However, four stations were found to have better performance when observing either Main Belt Asteroids (MBAs) or Near Earth Asteroids (NEAs).

The discussion on the statistic distribution of the residuals for *Steward Observatory, Kitt Peak-Spacewatch* is useful to investigate what causes the high gap between the RMSs for the different asteroids.

Future developments of this work could lead to the definition of different weights for the observations between NEAs and MBAs for the stations that show different measurement behaviours according to the asteroid class. This is likely to improve the whole OD pipeline. In this sense, one may think of the observations of Main Belters provided by the so-called *NEO Follow-up stations* that, as suggested by the name, are likely to be more specialized in providing accurate position for NEAs, leading to less accurate data for MBAs, as it has been shown for E12.

**References**

[1] Mario Carpino, Andrea Milani, and Steven R. Chesley. "Error statistics of asteroid optical astrometric observations". In: *Icarus* 166 (2 Dec. 2003), pp. 248–270.

[2] Steven R. Chesley, James Baer, and David G. Monet. "Treatment of star catalog biases in asteroid astrometric observations". In: *Icarus* 210 (1 Nov. 2010), pp. 158–181.

[3] D. Farnocchia et al. “Star catalog position and proper motion corrections in asteroid astrometry”. In: *Icarus* 245 (Jan. 2015), pp. 94–111. ISSN: 00191035.

[4] Siegfried Eggl et al. “Star catalog position and proper motion corrections in asteroid astrometry II: The Gaia era”. In: *Icarus* 339 (Mar. 2020).

[5] Peter Vereš et al. “Statistical analysis of astrometric errors for the most productive asteroid surveys”. In: *Icarus* 296 (Nov. 2017), pp. 1339–1351.

[6] Juan Cano et al. “Recent Evolutions in ESA’s NEO Coordination Centre System”. In: Apr. 2019.

[7] Richard Dekany et al. “The zwicky transient facility: Observing system”. In: *Publications of the Astronomical Society of the Pacific* 132 (1009 Mar. 2020).

[8] Richard Wainscoat. “The Pan-STARRS search for Near Earth Objects”. In: vol. 2016-June. IEEE Computer Society, June 2016.

2022-09-22

# Statistical analysis of a weighting scheme for asteroid observation astrometric errors taking into consideration the classification of the observed asteroids

Stronati, Nicolo'

International Astronautical Federation (IAF)