Searching for NEOs in SDSS Stripe 82

Andria C. Schwortz

Department of Physics and Astronomy, University of Wyoming, 1000 E. University Ave., Laramie, WY, 82071

aschwort@uwyo.edu

ABSTRACT

Near-Earth objects (NEOs) can pose a significant threat to life on Earth, depending upon their probability of impact, size, and how long ahead of a potential impact they are discovered. Mining SDSS data provides a nice opportunity for us to expand the number of NEOs identified without requiring additional equipment or time on existing telescopes. Herein I propose a method of mining SDSS DR8 coadd data and time dependent SDSS Stripe 82 data to discover fast-moving asteroids. This method results in the identification of an upper limit of 233 candidate NEOs within Stripe 82 posessing velocities in the range $0.5 < v < 2.5$ deg/day.

1. Introduction

The Sloan Digital Sky Survey (SDSS) is a multi-wavelength survey of approximately 35% of the sky, begun in the year 2000. Included in this survey is a region of the sky referred to as Stripe 82, covering an area measuring approximately 270 deg^2, or $RA = [315°, 59°]$ and $Dec = [-1.25°, 1.25°]$. Stripe 82 was repeatedly observed, with the intent of allowing astronomers a unique look at time dependent phenomena.

1.1. Statement of problem

In this project, the goal is to find asteroids within Stripe 82, specifically focusing on near-Earth objects (NEOs) with an angular speed of 0.5-2.5 deg/day.

1.2. Literature review

The main example of previous work on this topic is the SDSS Moving Object Catalog (SDSS-MOC). SDSS takes images in five wavelength bands (ugriz). Since these bands are taken sequentially in time, for a moving object the coordinates of the object will vary with band (and time). These are expressed in the tags "rowv" and "colv" (velocity across the rows and columns, respectively, of the CCD), in deg/day. The SDSS pipeline examines all starlike sources and determines a $v = \sqrt{rowv^2 + colv^2}$ for each. In DR1, the SDSSMOC contained all objects with $0.05 < v < 0.5$ deg/day (Ivezić et al. 2002). This was a total of $\sim 10,000$ asteroids at the time, with a prediction that by completion SDSS would identify 100,000 asteroids (Ivezić et al. 2001). I have not been able to identify any previous work using Stripe 82 data to find asteroids, and according to IAU Minor Planet Center (2013), there have been no dedicated surveys covering the entirety of the Stripe 82 region of the sky, so while there may have been individual asteroids discovered which cross this area of the sky, there does not appear to have been any systematic study.

2. Method

For this project I first searched for sources located within Stripe 82 with a velocity of $0.5 < v < 2.5$ deg/day according to DR8 rowv and colv values; see Sec. 2.1. I further culled this list by examining the Stripe 82 data; see Sec. 2.2. The overall routine is called "master".

2.1. Stage 1: DR8 catalog data retrieval and first cut

Within the area of the sky covered by Stripe 82, there are around 6 million sources which the SDSS pipeline has identified as being star-like (point sources). Due to their distance from Earth and small physical size, asteroids will be identified as point sources in SDSS data, so this list of 6 million sources must be culled to retain only those objects with a velocity in the target range.

This first stage was performed in "makefits". Sdss_sweep_data_index was used to find DR8 files containing sources in Stripe 82. I constructed circles to span the full RA/Dec range, spaced to cover Stripe 82, resulting in 43 of these circles (see "makecircles"). The circles were passed into sdss_sweep_data_index in "findfiles" and the names of appropriate .fits files retrieved along with the indices (minind and maxind). Any cases where *all of* the fits file name, minind, and maxind were identical were discarded. Lastly "buildsources" read in the files and kept sources inside[1] Stripe 82 and with appropriate velocities.

Running this program using the velocity limits as $v = [0.05, 0.5]$ deg/day as per Ivezić et al. (2002) found 1,388,579 candidate objects. As this number of candidates would take a prohibitive amount of time to analyze in Stage 2, I instead chose to focus on $v = [0.5, 2.5]$ deg/day.

---

[1]The circles must contain some area outside Stripe 82 for complete coverage. Each circle had a radius of 1.7676968° to ensure that it reached from a Dec of 0° to $\pm 1.25^\circ$ along a radius oriented at 45° to the RA/Dec axes, and accounting for the Dec dependence of RA. Radius was calculated via djs_diff_angle.

2.2. Stage 2: Stripe 82 data retrieval and second cut

"Pystripe" (called by "secondcut") stepped through each of the candidates found in Stage 1 and performed a SQL query (Myers 2013) to Stripe 82, returning all epochs of an object's RA, Dec, g-band psf magnitude, g-band MJD, and distance. A line was fit to the distance vs. MJD using mpfitfun, with the slope being the velocity which was then checked against the target velocities. The candidate objects' RA/Dec and velocity were then written out to the file "data/secondcut.fits".

3. Results and Discussion

From the 6 million stars in Stripe 82, Stage 1 culled these down to 42,678 candidates, and Stage 2 further culled these to 233. A histogram of the velocities of the 42,678 Stage 1 candidates is shown in Fig. 1. As expected, there are fewer candidates at higher velocities.

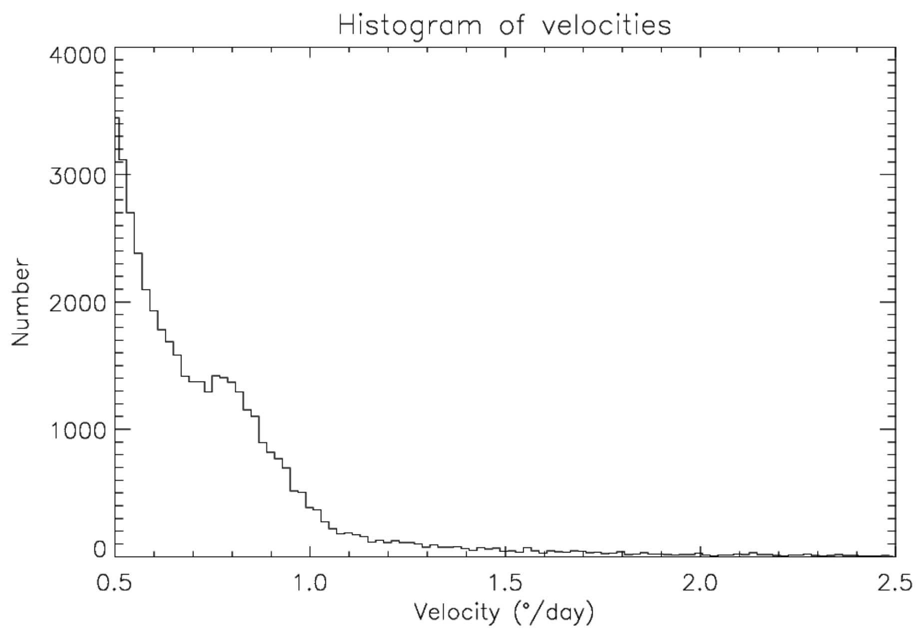

Fig. 1.— A histogram of the velocities of the 42,678 Stage 1 candidates, in 100 bins.

3.1. Flaws

Fig. 2 is a plot of a random candidate's distance (in degrees measured from the input RA/Dec) as a function of MJD (in days). It is clear that the best fit line found by mpfitfun is not correct. Varying the initial guess for slope and intercept (see line 65 in "pystripe") results in very different results for the output slope found by mpfitfun, see Fig. 3.

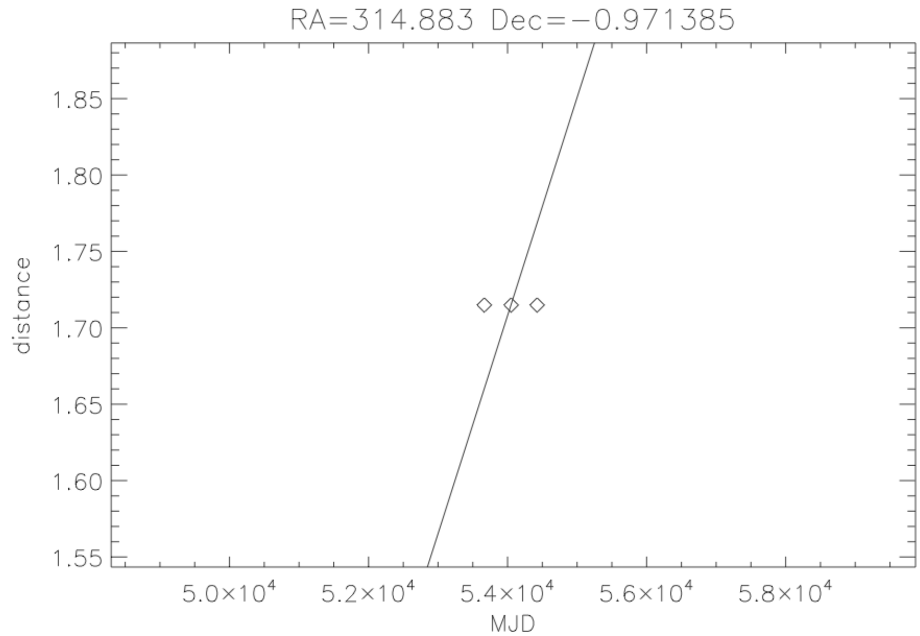

Fig. 2.— Stage 2 output: Plot of angular distance (in degrees) from input RA/Dec as a function of MJD (days) for one of the Stage 2 candidate NEOs. The solid line indicates the best fit line obtained by mpfitfun. It is clear that this line does not actually fit the data displayed.

Since the program is not finding the correct slope, the list of 233 candidates returned cannot be correct. Possible solutions are to (A) determine a way to force mpfitfun to come to a robust solution (e.g., using the Cash statistic which is supposed to be better for small numbers, or by forcing it to perform more iterations or with different step sizes), (B) eliminate sources with a small number of data points in Stripe 82, or (C) use a different routine for line fitting such as linfit.

The run time for the entire program is constrained by stripe82query.py. Assuming each query takes 1.5 s, the expected amount of time to repeat Stage 2 is $42,678 \times 1.5$ s $= 6.4 \times 10^4$ s $= 17.7$ hrs, so I am unable to re-run the entire program with these new criteria. However I tested a subsample of Stage 1 objects with changes (B) and (C) listed above; Table 1 summarizes the projected number of candidates as a result.

4. Future work

The full (19-hr) program should be rerun with the assorted versions of line fitting to produce a final count of candidate NEOs, and each source's RA/Dec outputted as a function of MJD and checked against the IAU/MPC catalog of confirmed asteroids. Data for confirmed asteroids should

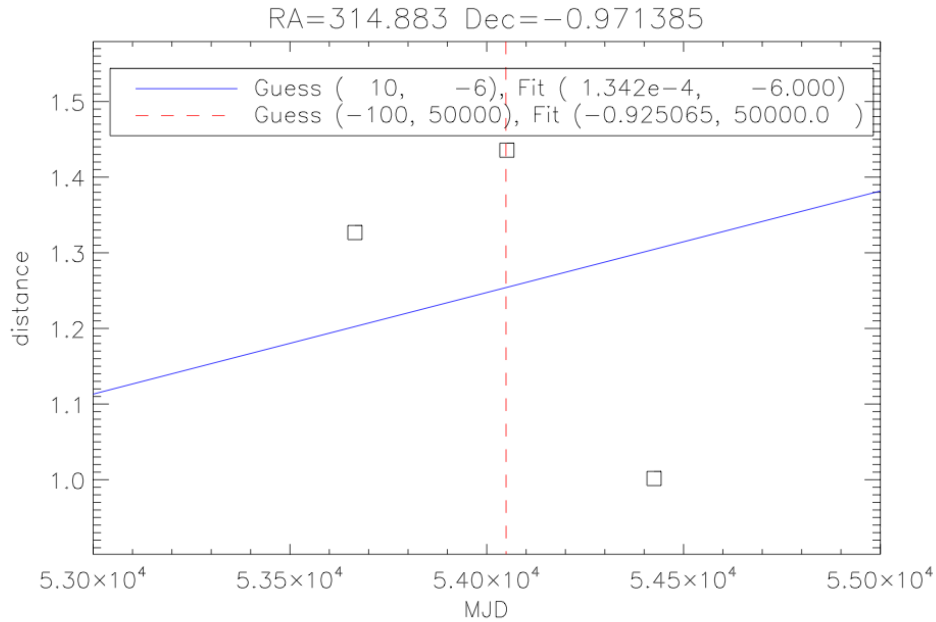

Fig. 3.— Stage 2 check: refitting the data for the source in Fig. 2. As shown by the lines, the slope and intercept found by mpfitfun was very sensitive to the initial guess for the slope and intercept, listed as (slope, intercept) in the legend.

be submitted to the MPC to improve accuracy. Data for new NEOs should also be submitted to the MPC for confirmation by other observers/methods. Should it be possible to perform this work on a supercomputer, the range of velocities studied could be expanded.

5. Conclusion

After running for more than 19 hours, the original version of my program returned 233 candidate NEOs with velocities in the target range, however this number is not accurate due to flawed line fitting. The original line fitting configuration found that 0.80% of a subsample of 500 sources fit the criteria for NEOs. Including more robust requirements for line fitting can be expected to reduce the number of candidate NEOs found within the Stripe 82, as is evidenced by other versions of the line fitting returning zero candidates. Therefore, I conclude that there are fewer than 233 candidate NEOs detectable by this method in SDSS Stripe 82 with velocities within the range of $0.5 < v < 2.5$ deg/day.

Fit method	Sample size	N_{points}	$\%_{cands}$	N_{pred}	N_{act}	Good fit?
mpfitfun	42,678	3	0.55	-	233	poor
mpfitfun	500	3	0.80	341	-	poor
mpfitfun	500	10	0.00	0	-	-
linfit	500	2	0.00	0	-	-

Table 1: Projected candidates for various line fitting methods. N_{points} is the number of data points required prior to fitting a line (mpfitfun requires 3 data points to run, linfit allows 2). Since no candidates were returned when using linfit with 2 data points, I did not repeat for 3 data points. $\%_{cands}$ is the percentage of objects examined which were returned as candidates. N_{pred} is the predicted number of candidates after running on all 42,678 Stage 1 sources. N_{act} is the actual number of candidate NEOs discovered when the full program was originally run.

REFERENCES

IAU Minor Planet Center. 2013, Sky Coverage Plots

Ivezić, Z., Juric, M., Lupton, R. H., Tabachnik, S., & Quinn, T. 2002, Astronomical Telescopes & Instrumentation, 98103

Ivezić, Z., Tabachnik, S., Rafikov, R., Lupton, R. H., Quinn, T., Hammergren, M., Eyer, L., Chu, J., Armstrong, J. C., Fan, X., & et al. 2001, The Astronomical Journal, 122, 27492784

Myers, A. 2013, stripe82query.py

This preprint was prepared with the AAS LaTeX macros v5.2.