Tests of Detection Density and Filtering Conditions of WMOPS - 16 May 2013

Joe Masiero and the WMOPS team

1. Regions with high number density of detections

Areas on the sky containing dense nebulosity at W3 wavelengths result in unusually large numbers of detections per frame, particularly as the SNR cut is lowered. As these detections do not correspond to static point sources and are more akin to amplified background fluctuations, they are not eliminated by comparisons with the Atlas images or stationary object checks. Nebulous regions tend to have a large enough extent that linking of noise sources into false tracklets becomes frequent, which creates a large number of false tracklets to QA and slows down glob processing considerably.

By restricting the number of detections per frame that are ingested (post-SOR) this effect can be mitigated. Figure 1 shows the distribution of post-SOR detections per frame in a non-nebulous region of the Cryo data for a variety of SNR cutoffs. Figure 2 shows the median number of detections per frame as a function of SNR, as well as a rough power-law fit to the data of $N = 3370\ \mathrm{SNR}^{-3.05}$. An upper limit on number of detections per frame of $N \sim 200$ will encompass the majority of nominal frames, even down to SNR= 3. For problem frames, increasing the SNR cutoff until the total number of post-SOR detections is below this level should mitigate many of the problems associated with false tracklets being built out of nebulosity-detections.

2. SNR offset from cryo to post cryo

Due to changes implemented in the post-cryo determination of background noise, in regions of high stellar density SNR will differ for otherwise identical sources when compared to the Cryo dataset. This effect becomes more prominent as the region in question approaches the galactic plane. This is illustrated in Figures 3 and 4, which show the SNR measured for stationary objects in a region of sky observed during both the Cryo and Post-Cryo missions for galactic latitudes of $b = -2^\circ$ and $b = -22^\circ$ respectively. Galactic latitude for both regions is $l = 60^\circ$ and the red line indicates a running mean of all detections. In extreme cases near the galactic plane a detection of SNR= 4 in the post-cryo dataset corresponds to a cryo detection of SNR$\sim$ 2.9. This, combined with the increase in background source density in W2, results in a large increase in number of detections per frame and thus tracklets per run for post cryo runs.

Figure 5 shows the equivalent Cryo SNR value for four Post cryo detection levels as a function of galactic latitude for three tested longitudes. The overall trend is for the SNR levels to be approximately equal at latitudes above $|b| \sim 20$ but approximately one point lower at the regions closest to the plane, however these is significant variation based the specific galactic longitude,

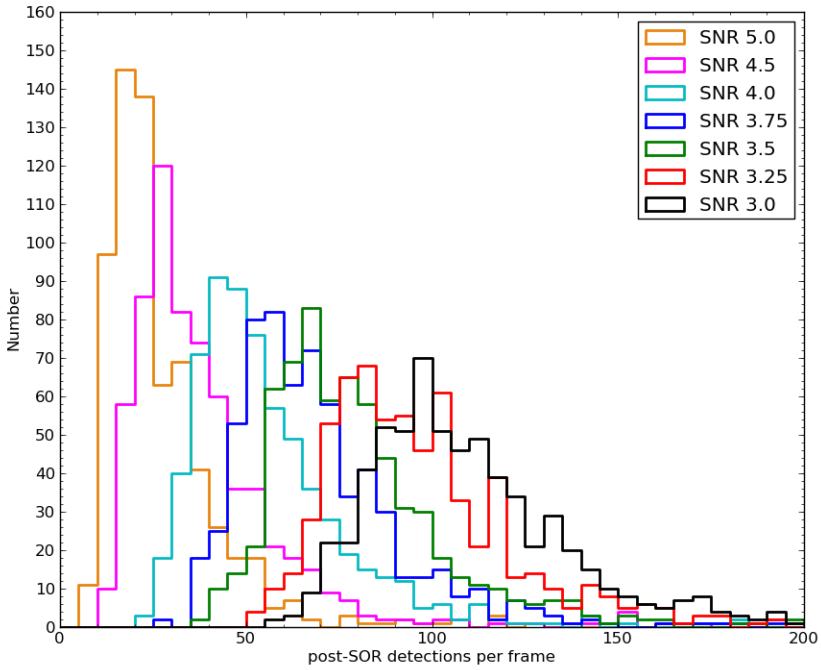

Fig. 1.— Detections per frame for various SNR cuts

likely reflecting the effect of local stellar density and nebulosity.

3. Tests of filtering options to reduce false tracklet load

3.1. Ecliptic Latitude

At very high ecliptic latitudes overlapping coverage from scans over long periods of time results in a rapid increase in the number of possible combinations of detections into potential tracklets. This is particularly true in cases where five detections no longer represents approximately half of the images of a particular region, but a significantly smaller fraction of them.

Figure 6 shows the distributions of all tracklets produced for QA for the 13120a run, as a function of both ecliptic latitude (both number and number density) and tracklet length. The majority of tracklets with short lengths fall in the highest ecliptic latitude bin ($85° - 90°$), and when weighted by steradian area the tracklet number density is $1 - 3$ orders of magnitude higher than regions closer to the ecliptic.

This is particularly apparent in Figure 7 which shows a scatter plot of ecliptic latitude compared

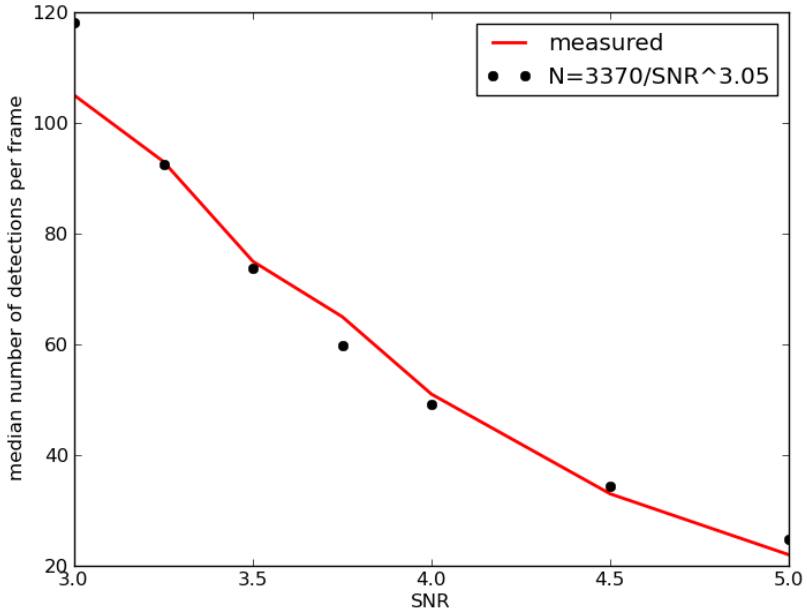

Fig. 2.— Characteristic number of detections per frame compared to SNR

with tracklet lengths. Length values have been modified by a ± 0.4 random jitter for clarity. The numbers overlaying the areas show the number of points in the given bin of width 10° for track length values of $5-7$ or at all latitudes for tracklets of length $8-9$. Horizontal over-densities in this plot are likely artifacts of nebulous or otherwise confused regions. While rejecting short tracks in the ecliptic polar regions will greatly reduce the number of tracklets, the QA load is still strongly inflated by these nebulosity tracklets. It is also clearly apparent that tracklets of length 5 are still quite important for the auto-accepted population.

Figure 8 shows the differentiation between the ecliptic coordinates of tracklets that were auto accepted by SSOID, rejected by human QA, and accepted by human QA. Note that QA was only done on a subset of the tracklets produced. Not surprisingly, the majority of SSOIDed objects fall at ecliptic latitudes within $\pm 30^\circ$. In this plot, the QA rejections show strong clumping effects, which is a result of both the high density of false tracklets in regions of nebulosity or high ecliptic latitude, as well as the non-random ordering of tracklets in the QA page (a function of glob writeouts). This also helps explain the larger longitudinal extent of the auto-accepted tracklets, as the tracklets on the QA page tend to be ordered in time.

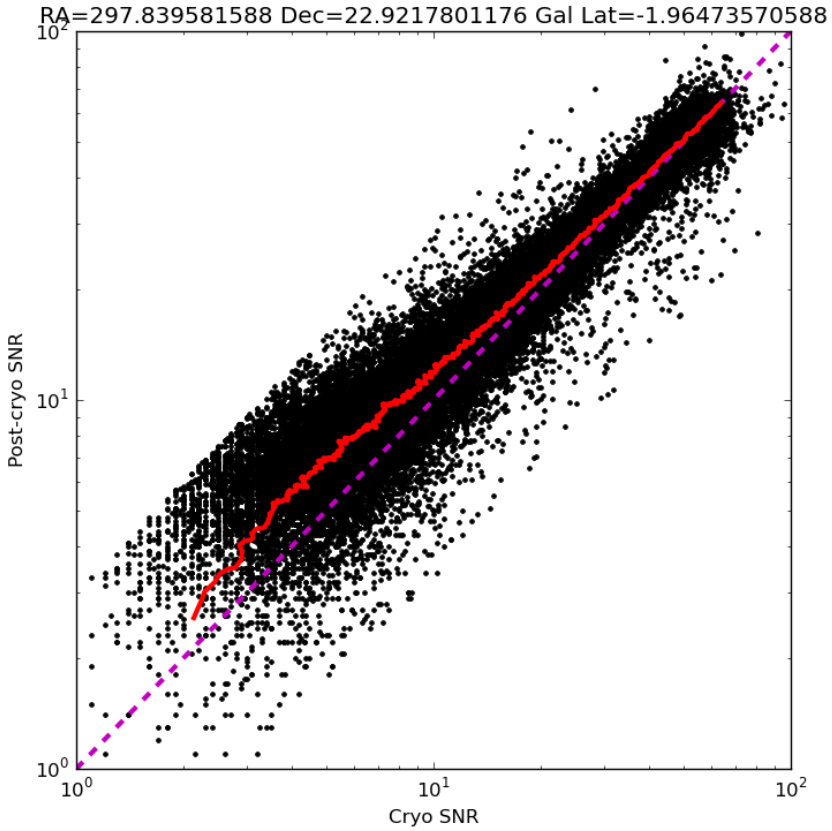

Fig. 3.— Comparison of cryo and post-cryo SNRs at $b = -02$

3.2. Digest score filtering

Each tracklet identified by the WMOPS software is run through the Minor Planet Center's Digest scoring algorithm to determine the likelihood that the linked detections represent a real, physically-possible orbit. Thus, the Digest score has the potential to be used as an easy filter of obvious false-linkages. The concern with this filtering is that tracks from real objects that include a single bad or blended detection may also return bad Digest scores and an overly aggressive filter will reject real objects.

Figure 9 shows distributions of Digest scores for all auto-accepted SSOIDed tracklets, the tracklets rejected by human QA, and the tracklets accepted by human QA. The scores of the auto-accepted tracklets (assumed to represent well what the score of a real previously unknown object would be) group around a score of $d \sim 1$, with a secondary grouping around $d \sim 4$. Conversely, human-rejected tracklets peak around $d \sim 7$ and a long tail at higher scores. This seems to be

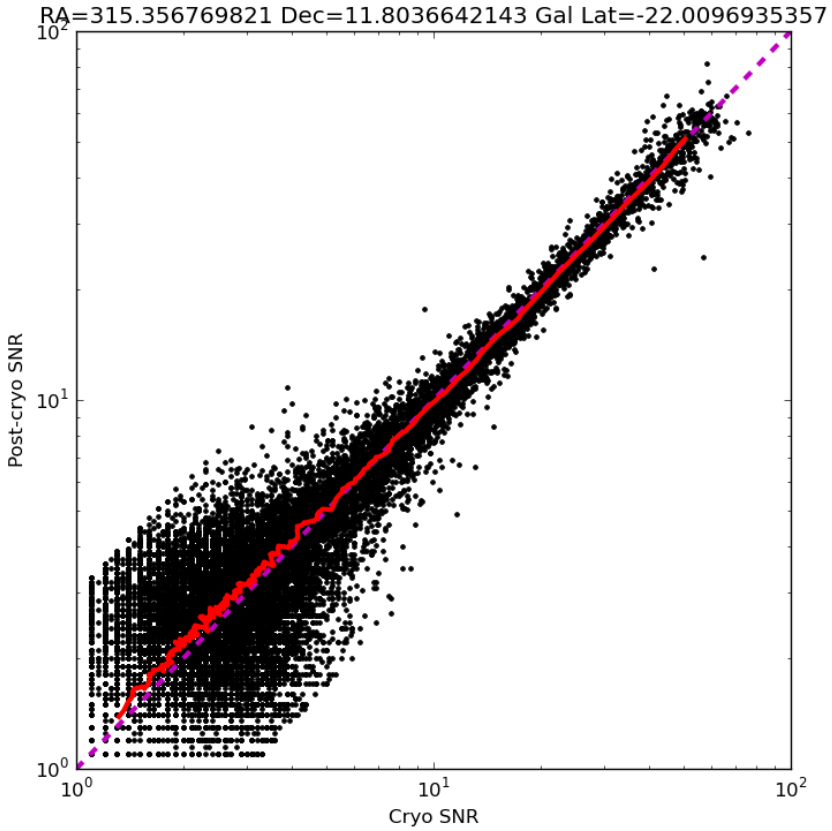

Fig. 4.— Comparison of cryo and post-cryo SNRs at $b = -22$

comparable to the distribution of human accepted tracklets as well, though this group suffers from very small number statistics. Only a 6 auto-accepted tracks ($\sim 0.06\%$) had $d > 20$ and all of these tracks were linkages where a single detection (usually the last) was a false-link. The one human-accepted track with $d > 20$ upon further review was not a real object, a fact that would have been identified in the cleaning step prior to submission to the MPC. A requirement of $d < 30$ would thus encompass all real objects identified in the data.

Figure 10 shows the expected impact of this cut on Digest score. The majority of rejected tracks would be at ecliptic latitudes between $55° < |l| < 80°$. This cut would compliment an ecliptic latitude based SNR- or track-length filter as these high-Digest tracklets are longer than other tracks at those latitudes. This high-Digest population represents $20 - 25\%$ of the human-rejected population and thus would be expected to reduce the amount of QA required by a comparable fraction.

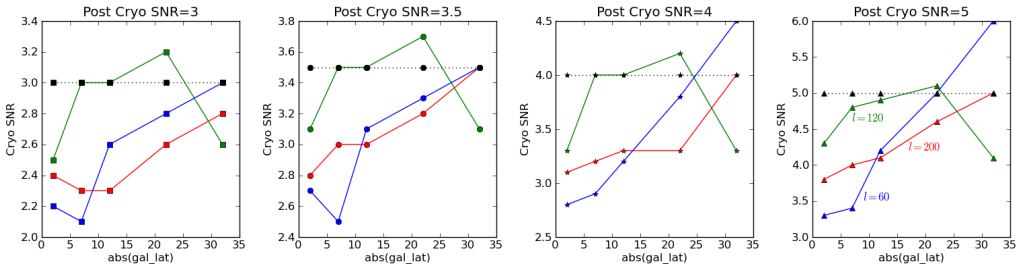

Fig. 5.— SNR vs galactic lat for various post-cryo SNR cuts

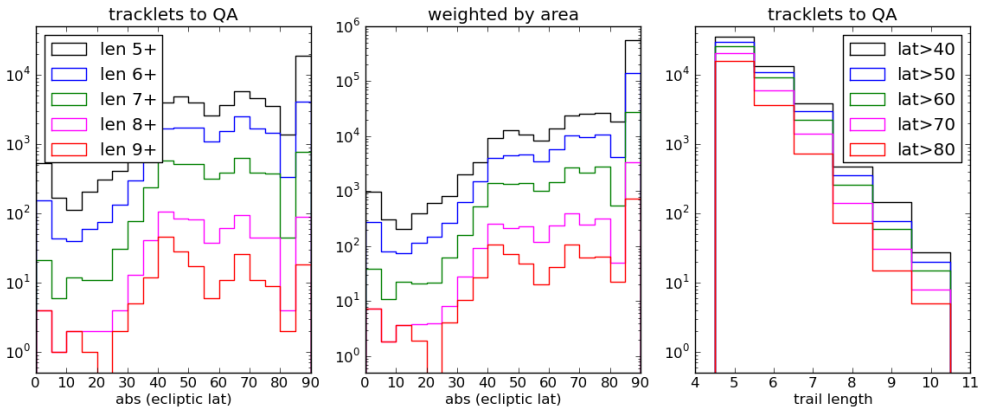

Fig. 6.— Distributions of tracklets as a function of ecliptic latitude and track length

However, this picture is somewhat complicated when tracks are differentiated by ecliptic longitude and latitude zones. Figure 11 shows the digest scores for rejected tracks for all latitudes separated into the two ecliptic longitude strips covered during this sample run. The peaks in the digest scores distributions correspond to the location of the equatorial poles in the ecliptic coordinate space. Possible physical or software causes for this effect are being investigated.

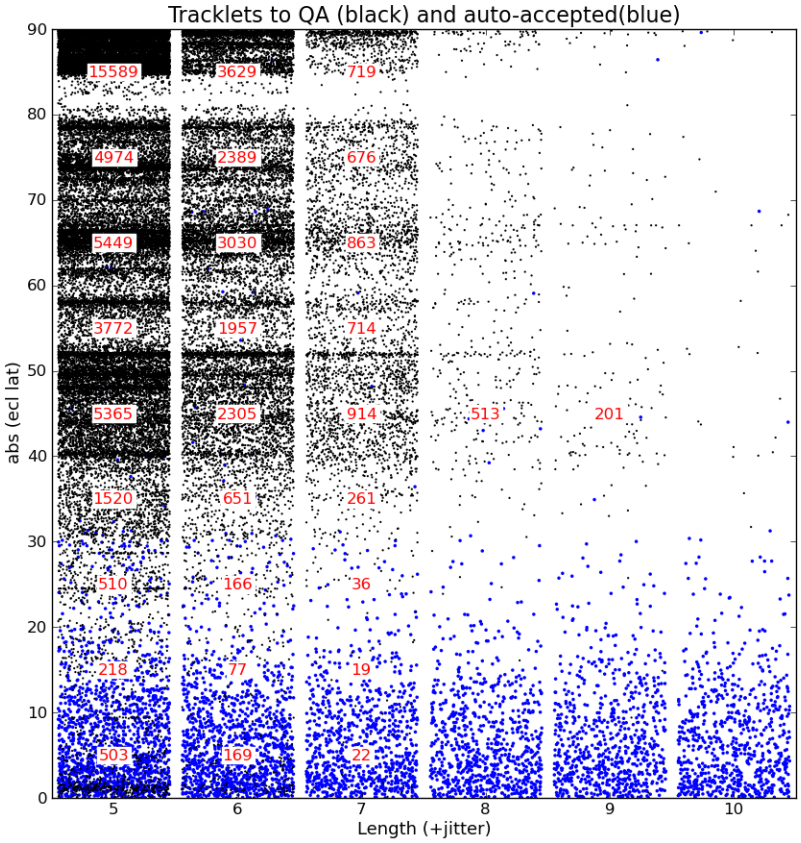

Fig. 7.— Comparisons of Ecliptic Lat and track length

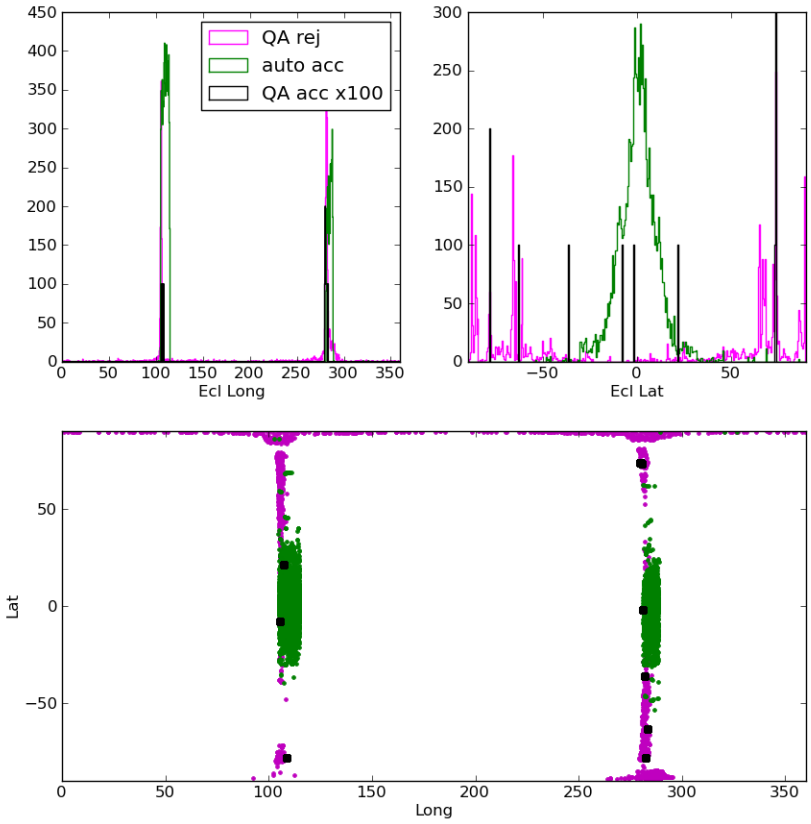

Fig. 8.— Ecliptic Lat and Long distributions for rejected and accepted tracklets

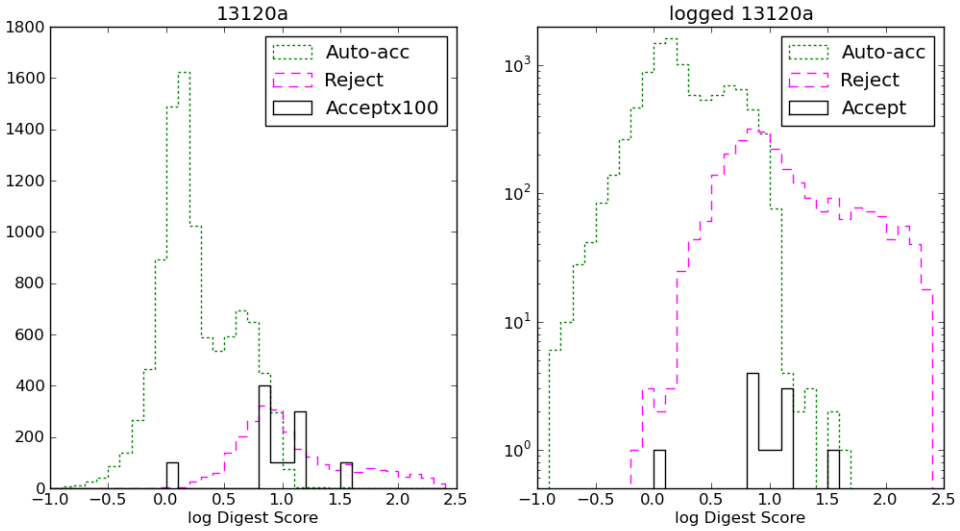

Fig. 9.— Digest score distributions for SSOID accepted (green), human QA rejected (magenta) and human QA accepted (black) tracklets. The left plot shows bin-heights in linear-space, with QA accepted bins magnified by a factor of 100, while the right plot shows the bin heights in log-space.

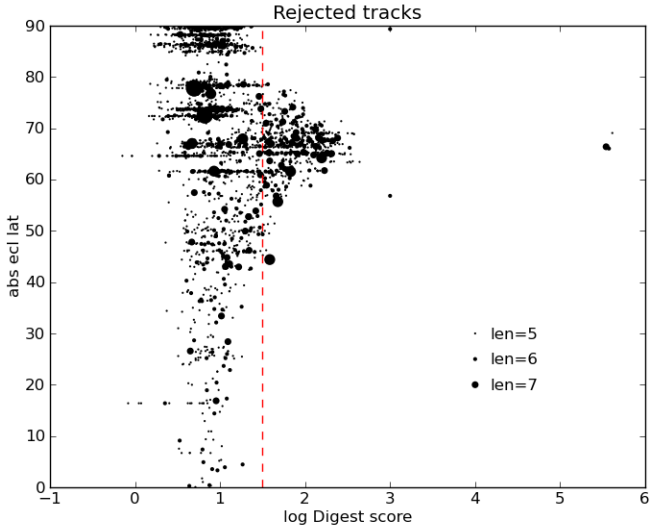

Fig. 10.— Ecliptic latitude vs Digest score for all human-rejected tracklets, with track length indicated by point size. Tracks to the right of the vertical red dashed line would be removed by the proposed digest filter. Horizontal striations are a result of the clumping discussed above.

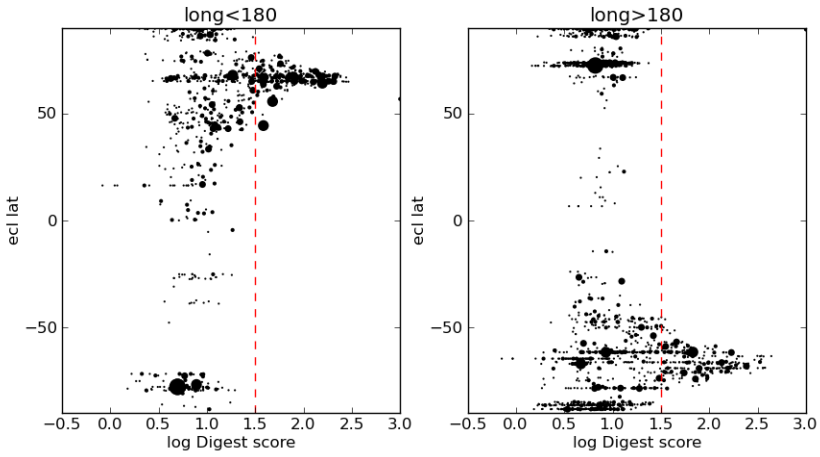

Fig. 11.— The same as 10 but with ecliptic longitude and latitude zones separated. Digest scores peak near the equatorial poles.

3.3. Changing SNR threshold from median value to minimum value

For the majority of the test runs described here, a reliability filter was introduced that required low SNR tracklets ($SNR < 5$) to have length 6 or greater to be passed to human QA. Brighter tracklets were still allowed to pass with length 5. An extension of this based on the ecliptic latitude filtering results from § 3.1 can be implemented to further reduce the QA load, by required length 7 for tracklets with $SNR < 5$, length 6 for tracklets with $SNR < 6$ and length 5 for all brighter tracklets. However even with this filter in place, a large number of length 5 tracklets were being passed to QA despite appearances that the detections were not of high enough significance.

The cause of this was traced to low level cosmic ray hits that were being identified of moderately high SNR detections, which resulted in the median SNR being much higher than would be expected. To mitigate this problem, it is possible to instead set SNR thresholds against the minimum SNR detection of the tracklet. Figure 12 shows the distribution of minimum SNR values for tracklets that were auto-accepted, rejected in human QA, or accepted by human QA. As expected, all distributions show a peak around $SNR = 3.75$ which was the cutoff for the run (implying a track composed mostly of detections near the limit). However there are long tails of tracks at low minimum SNRs that indicate tracklets that are comprised mostly of noise points. Figure 13 highlights the relation between minimum SNR and tracklet length for all three of these groups of tracklets. The majority of false tracklets have short lengths and low minimum SNRs, meaning a stepped filter in minimum SNR would result in a large improvement of the tracklet QA load.

Figure 14 shows the distributions of real and rejected tracklets broken out by length bin. Very few false tracklets have lengths of $l \geq 7$. Length 5 tracklets are split approximately evenly between real and rejected tracklets in this plot, however the auto-accepted tracklets cover the full run, while the human QA only processed $\sim 5\%$ of all tracklets to QA. Thus at length 5, false tracklets outnumber real tracklets by at least $20:1$. We note that these distributions are after median SNR filtering, so the true false-to-good ratio is likely significantly larger. A cut of $SNR_{min} = 5$ at length 5 would eliminate the majority of bad tracklets, and thus a large fraction of the QA load, with only minimal losses to any potential real objects in the data. It is also probable that some of the auto-accepted tracklets, especially those at short lengths and low SNR_{min}, are duplicate tracklets that are represented by longer or higher SNR tracklets elsewhere in the run.

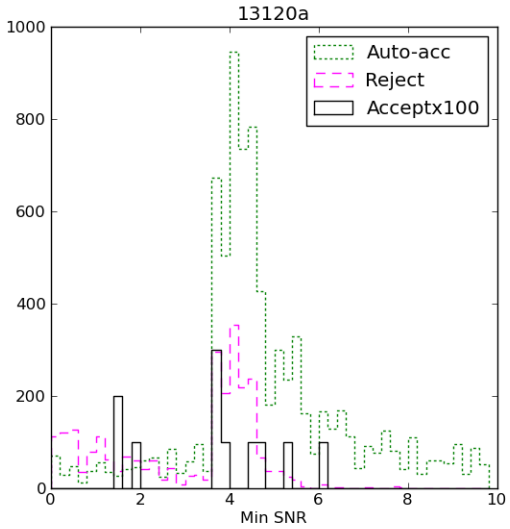

Fig. 12.— Distribution of minimum SNR values for auto-accepted, human-rejected, and human-accepted tracklets. Tracklets with SNR below the run limit of 3.75 are likely non-detection identified by the extraction pipeline. Human-accepted distributions are magnified by a factor of 100.

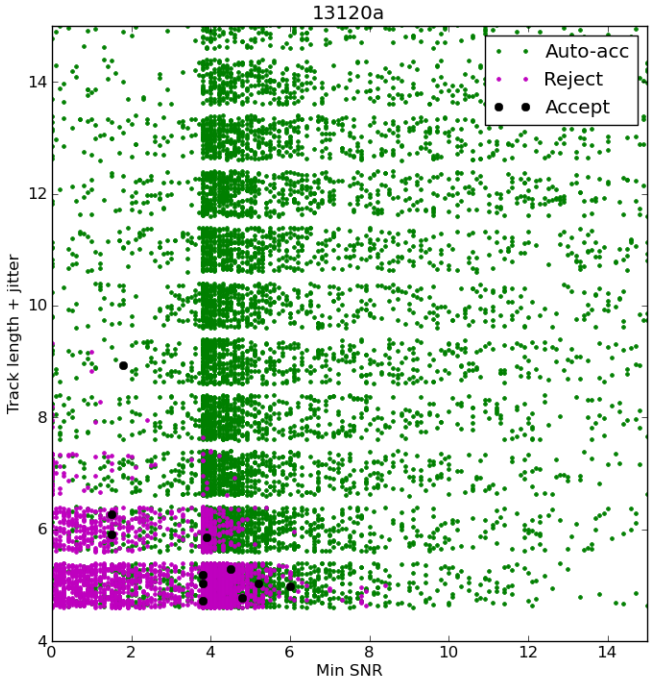

Fig. 13.— Track length vs minimum SNR for all auto-accepted, human-rejected and human-accepted tracklets. Track length has been jittered by a random number in the range of ±0.4 for clarity. The majority of bad tracklets have short lengths and low minimum SNR.

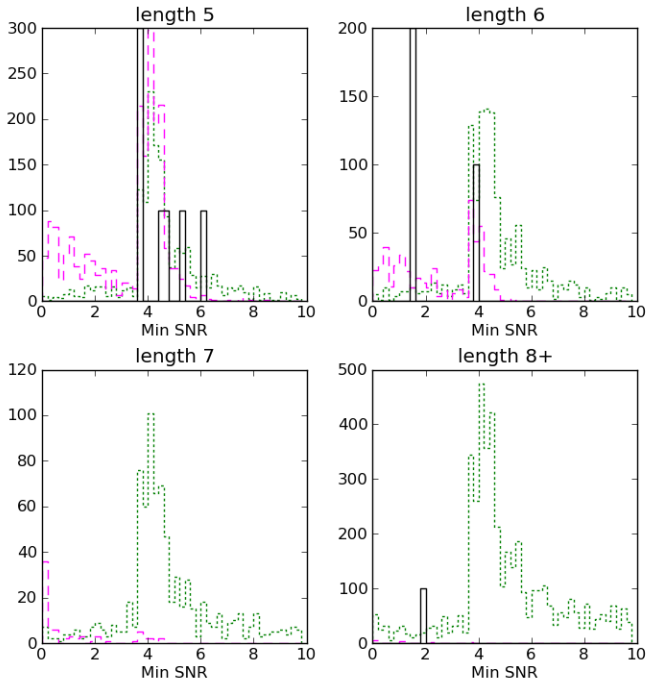

Fig. 14.— The same as Figure 12, but separated by tracklet length.

3.4. Filtering by the linearity of the declination vs time motion

WMOPS links tracklets by looking for clusters of objects with similar on-sky motions, velocity positional angles, and positions on the sky. However, right ascension and declination are treated as rectilinear coordinates for these calculations. For objects near the equatorial poles, this results in calculated right ascension motions that are anomalously large when compared with the declination motions or true, on-sky motions. The result of this discrepancy is that when total motion is calculated it is dominated by the artificial right ascension rates. This results tuples being linked into tracklets that have non-linear and non-physical changes in declination. Source density increases when the SNR cutoff is lowered, and as coverage at high ecliptic latitudes increases when larger data sets are considered, which results in more objects being linked with these non-linearities.

Although the optimal solution is to revise the linking algorithm to properly handle motions at high declinations, any rectilinear re-projection will suffer from the same problems and a true spherical consideration may result in a large increase in run time. An alternate solution, and a method of correcting previously run QA sets, is to filter identified tracklets based on the linearity of the declination motion before sending these for human QA.

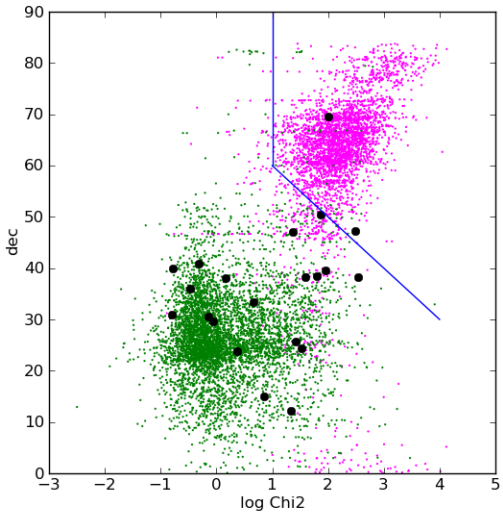

Fig. 15.— Declination vs reduced χ^2 for auto-accepted (green), human-accepted (black) and human rejected (magenta) tracklets. The blue line indicates the proposed region to reject.

In Figure 15 we show the reduced chi-squared from a linear fit to declination versus time for accepted and rejected tracklets from the 13155b run. The overall trend is for linearity to get significantly worse at higher declinations, and for this region to be populated with a large number of bad tracklets. By setting a boundary that separates the majority of the bad tracklets from the worst real tracklets (nicknamed a 'Nevada cut') we can auto reject the majority of bad tracklets and

thereby reduce the human QA load. We note that there are some object identified as real within this boundary, however these are either long tracklets that do show real curvature in their on-sky motion, and thus fail linearity tests, or blends of majority real detections with a few particularly bad links to noise points that result in high χ^2 values. To reduce rejections due to the first case, we do not auto-reject tracklets with 15 or more detections. The second problem however is also illustrated in the auto-accepted tracklet χ^2 values.

There is a clear extension to higher χ^2 values in the distribution of good tracklets with respect to χ^2, which represents real objects where a single bad detection (usually a cosmic ray or other noise point) had been incorrectly linked into the track. In Figure 16 we show a reanalysis, now rejecting from every tracklet the one worst-fitting point and then re-computing the χ^2. This results in the χ^2 values for all tracklets decreasing, as we would expect from rejecting the worst point. Notably, the extension in the good tracklets at larger χ^2 is largely removed, and most of the real objects inside the Nevada cut also move beyond the bounding line. Using the outlier-rejected linearity test, and the Nevada cut, we can automatically reduce the number of tracklets for human QA by over 2/3, with minimal-to-no loss of real objects, and only a small increase in WMOPS run time.

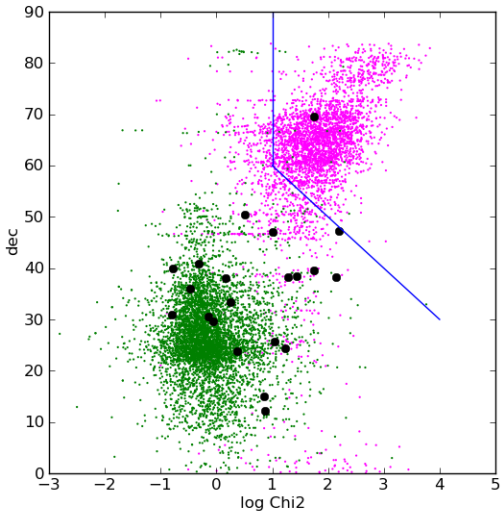

Fig. 16.— The same as Figure 15, but where one worst detection is rejected from the χ^2 analysis.

Figure 17 shows a comparison of declination and linearity after the application of the "Nevada" filter for a single hemisphere run over 3.75 days. For the ease of implementation, the filter boundaries have been discretized, with a preference to a less-restrictive cut. Some tracklets remain present in the QA list despite being within the filtered region; these tracklets either have track lengths of 15 or greater, or were pre-approved by WMOPS as SSOID-identified objects, as discussed above. This latter case also explains the linear features in declination-space: these are a series of tracklets

composed of 4 good detections with a variety of different bad fifth detections, that will have the same declination but different χ^2 values.

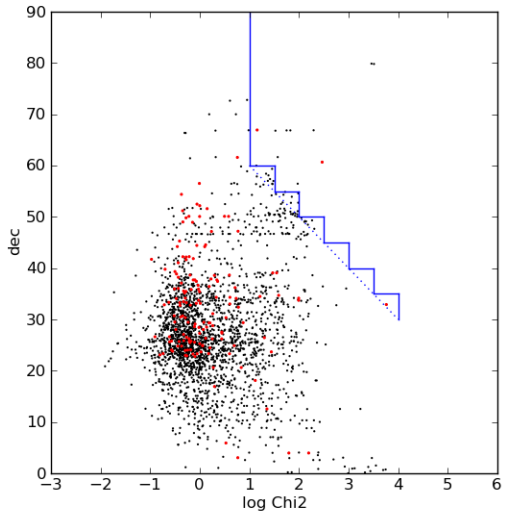

Fig. 17.— The same as Figure 15 for a pre-QA dataset with a discretized "Nevada" filter applied during QA product generation (solid blue line). The dotted blue line shows the idealized filter case. Red points indicate those with more than 15 detections.