Detecting Active Comets with SDSS

Michael Solontoi [a], Željko Ivezić [a], Andrew A. West [b],
Mark Claire [a], Mario Jurić [c] Andrew Becker [a], Lynne Jones [a],
Patrick B. Hall [d], Steve Kent [e] Robert H. Lupton [c],
Tom Quinn [a], Gillian R. Knapp [c], James E. Gunn [c],
Craig Loomis [c]

[a] *University of Washington, Dept. of Astronomy, Box 351580, Seattle, WA 98195, USA*

[b] *MIT Kavli Institute for Astrophysics and Space Research, 77 Massachusetts Ave, Cambridge, MA 02139-4307, USA*

[c] *Princeton University Observatory, Princeton, NJ 08544, USA*

[d] *Department of Physics & Astronomy, York University, 4700 Keele St., Toronto, ON M3J 1P3, Canada*

[e] *Fermi National Accelerator Laboratory, Batavia, IL 60510, USA*

Number of pages: 28

Number of tables: 5

Number of figures: 19

Preprint submitted to Icarus *1 December 2008*

Proposed Running Head:

Detecting Active Comets with SDSS

Please send Editorial Correspondence to:

Michael Solontoi

University of Washington

Department of Astronomy

Box 351580, Seattle, WA 98125

Email: solontoi@astro.washington.edu

Phone: (206) 543-9095

Fax: (206) 685-0403

ABSTRACT

Using a sample of serendipitously discovered active comets in the Sloan Digital Sky Survey (SDSS), we develop well-controlled selection criteria for greatly reducing the number comet candidates selected from SDSS catalogs. After the follow-up visual inspection that rejected remaining false positives, the total sample of SDSS comets presented here includes 19 objects including their ecliptic latitude distribution, and the comet distribution in the SDSS color space. The good understanding of selection effects allow us to study population statistics, and we estimate the magnitude distribution down to $r \sim 18$. Perhaps the most surprising results are the extremely narrow color distribution for comets in our sample (e.g. root-mean-square scatter of only $\sim$0.06 mag for the $g - r$ color), and the similarity of comet colors to those of Jovian Trojans. We discuss the relevance of our results for upcoming deep multi-epoch optical surveys such as the Dark Energy Survey, Pan-STARRS and the Large Synoptic Survey Telescope (LSST), and estimate that LSST may produce a sample of about 10,000 comets over its 10-year long survey.

Keywords: COMETS, PHOTOMETRY

1 Introduction

The small bodies of the solar system offer a unique insight into its early stages and evolution. Only a few locations still exist where samples of the original solid materials of the solar nebula remain: The main asteroid belt, the Trojan populations of the giant planets, the Kuiper Belt just outside the orbit of Neptune, and far beyond the planets the objects in the Oort cloud. Today these populations can be studied as asteroids and comets. Understanding these populations, both physically and in their number and size distribution, is a key element in testing various theories of solar system formation and the evolution of our planetary system. Comets represent samples of the two farthest regions of the solar system, the Kuiper Belt and the Oort Cloud, regions whose great distances make in-situ studies difficult, and in the case of the Oort Cloud, beyond our current observational capabilities.

Comets are defined as objects that have activity: the ejection by out-gassing of small amounts of dust and gas, typically at small heliocentric distances. These materials interact with the solar wind and radiation, leading to the production of cometary coma and tails. Dynamically, comets tend to belong to one of two main groups: the Jupiter Family Comets (JFC), consisting generally of low inclination, prograde orbits, thought to originate in the reservoir of the Kuiper Belt, and Long Period Comets (LPC), whose orbits include a wide range of inclinations and directions, with aphelia far beyond the planets of the solar system. A common method used to separate the JFC, LPC and asteroidal type of orbits is the use of the Tisserand parameter with respect to Jupiter (Levison & Duncan, 1997) . This requires a comet to have a well characterized orbit, so it does not apply to comets with only single epoch observations, nor

can it distinguish the comets residing within the main asteroid belt (Hsieh & Jewitt, 2006) from the population of main belt asteroids.

Most photometric surveys of comets to date have targeted select groups of known comets, often only in one or two filters, to look for activity levels at various heliocentric distances, and to make estimates of their physical properties (Lowry et al., 1999, and references therein). Other surveys of known comets have focused on spectroscopy, or the use of narrow-band filters, in order to estimate physical properties and investigate the gas species (A'Hearn et al., 1995, and references therein). While these targeted surveys make great contributions to cometary knowledge, they do not offer the opportunity to study an unbiased selection of comets, or to be able to answer questions about the number of comets on the sky at a given time, and their distributions in magnitude, heliocentric distance and angular size.

As opposed to these targeted surveys, all-sky surveys such as the Lincoln Near-Earth Asteroid Program (LINEAR) detect vast numbers of solar system objects including comets (Stokes et al., 2000). These detections and cometary discoveries build up excellent catalogs of object populations for study, but do not have the ability to observe the comets in multiple filters.

The Sloan Digital Sky Survey (SDSS) (York et al., 2000, and references therein) offers an opportunity to observe a large number of comets in multiple filters. By the fifth public data release (DR5) almost 8000 square degrees of sky has been imaged (Adelman-McCarthy et al., 2007), thus allowing detection of both known, and as of yet, unknown comets on the sky for population studies and providing accurately calibrated five-color photometry. These data can be used for determining physical properties of the comets imaged by the survey. The

SDSS also allows direct comparison with other populations of small bodies (NEOs, Main Belt Asteroids, Trojans) imaged by the survey, all having been observed by the same instrument and calibrated to the same standards.

While specifically looking for known comets in SDSS data could produce a list of comets for detailed study, the ability to "mine" SDSS for comets is a way to turn it into a discovery survey as well. A method needs to be developed that can cull detections of comets from the vast amounts of SDSS data. Such a method will allow both known and undiscovered comets to be found in SDSS, and will also inform methods for detecting new comets in the upcoming next generation of large area surveys (e.g. Pan-STARRS and LSST, Kaiser et al., 2002; Ivezić et al., 2008)

Active comets are not easily identified in the SDSS imaging catalogs. Given the sky coverage, and the numbers of known comets on the sky, it is expected that a few dozen comets may have been imaged by the survey. Selecting these objects from the 215 million object DR5 catalog without a significant contribution by false positives is not trivial. Active comets are resolved objects, and have a measured angular velocity. However, a selection criteria designed to be sensitive to comets out to Jupiter's orbit also returns about 1% of the SDSS galaxies. Given the large number of galaxy detections ($\sim 10^8$), this selection results in about 1 million false positives. While most comets can be efficiently recognized in SDSS images via visual inspection, a sample of ~ 1 million candidates is too large to visually examine. These false positives could be easily rejected if a "second epoch" of images were available, but over most of the survey area the SDSS obtained only one epoch of data. As an alternative, we develop methods based on the use of SDSS galaxy spectroscopy as "second epoch" data as well as restricted search criteria in velocity and color

space, that can be used to prune the candidate list down to a number that is manageable for visual inspection.

In Section 2. we present and test selection criteria for identifying objects from the Sloan Digital Sky Survey catalogs that may be active comets. The analysis of the resulting sample of 19 confirmed comets, including the sample magnitude, sky and color distributions, is described in Section 3, and we discuss our results in Section 4. In the remainder of this paper, which focuses on selection methods, we use cataloged parameters automatically computed by the SDSS processing pipelines. In a sequel to this paper, we will present an analysis of SDSS data based on custom photometry of the individual observations.

2 Methodology

2.1 An Overview of SDSS

The SDSS is a digital photometric and spectroscopic survey which will cover about one quarter of the Celestial Sphere in the North Galactic cap, and produce a smaller ($\sim$ 300 deg^2) but much deeper survey in the Southern Galactic hemisphere (Stoughton et al., 2002; Abazajian et al., 2003, 2004, 2005; Adelman-McCarthy et al., 2006). SDSS is using a dedicated 2.5m telescope (Gunn et al., 2006) to provide homogeneous and deep (r $<$ 22.5) photometry in five band-passes (Fukugita et al., 1996; Gunn et al., 1998; Smith et al., 2002; Hogg et al., 2001; Tucker et al., 2006) repeatable to 0.02 mag (root-mean-square scatter, hereafter rms, for sources not limited by photon statistics, Ivezić et al., 2003) and with a zero-point uncertainty of $\sim$0.02-0.03 mag (Ivezić et al., 2004). The flux densities of detected objects are measured

almost simultaneously in five bands (*u, g, r, i*, and *z*) with effective wavelengths of 3540 Å, 4760 Å, 6280 Å, 7690 Å, and 9250 Å. The large survey sky coverage will result in photometric measurements for well over 100 million stars and a similar number of galaxies [1]. The completeness of SDSS catalogs for point sources is ∼99.3% at the bright end and drops to 95% at magnitudes of 22.1, 22.4, 22.1, 21.2, and 20.3 in *u, g, r, i*, and *z*, respectively. Astrometric positions are accurate to better than 0.1 arc-second per coordinate (rms) for sources with r $<$ 20.5 (Pier et al., 2003), and the morphological information from images allows reliable star-galaxy separation to r $\sim$ 21.5 (Lupton et al., 2002; Scranton et al., 2002). Analysis of SDSS galaxies show that the sky is accurately and robustly subtracted by the photometric pipeline. A compendium of other technical details about SDSS can be found on the SDSS web site (http://www.sdss.org), which also provides interface for the public data access.

The samples presented in this paper are based on the SDSS Fifth Public Data Release, hereafter DR5 (Adelman-McCarthy et al., 2007). More information about this data release can be found at http://www.sdss.org/DR5.

2.2 Moving Objects in SDSS

The SDSS, while mainly designed for observations of extragalactic objects, is significantly contributing to studies of the solar system, notably in the success it has had with asteroid detections, cataloged in the SDSS Moving

[1] The fifth public data release (DR5) covers almost 8000 deg^2 of sky, and includes catalog of 215 million objects (Adelman-McCarthy et al., 2007); see http://www.sdss.org.dr5/.

Object Catalog (hereafter SDSS MOC, Ivezić et al., 2001). This public, value added catalog of SDSS asteroid observations contains, as of its fourth release, measurements of 471,000 moving objects, 220,000 of which were matched to 104,000 known asteroids from the ASTORB file [2]. The quality of SDSS MOC data was discussed in detail in Ivezić et al. (2001) and Jurić et al. (2002).

The SDSS camera (Gunn et al., 1998) uses a drift-scanning technique and detects objects in the order *r, i, u, z, g*, with detection in two successive bands separated in time by 72 seconds. Moving objects appear to have their colors separated when color composite images are made (different bands are registered to the same coordinate system using stationary stars), and if moving fast enough, appear as streaks of individual colors. Asteroids in SDSS typically appear as unresolved (star-like) sources (a sequence of dots or streaks, see Figure 1). Comets are relatively easy to visually distinguish from other sources. An active comet is typically a resolved source, with color separation visible in the near nuclear coma, and is surrounded by a diffuse coma and sometimes tail(s). Figure 1 shows examples of a comet and an asteroid imaged by SDSS.

2.3 SDSS Magnitudes and Processing Flags

Unless otherwise specified, the magnitudes we present for the comets here are so-called model magnitudes. These model magnitudes are designed for galaxy photometry, and are determined by accepting the better of a de Vaucouileurs and an exponential profile of arbitrary size and orientation. The surface brightness profile of a comet differs from that of a galaxy, and thus these model magnitudes may not be optimal for comet studies. We will present analysis based

[2] see ftp://ftp.lowell.edu/pub/elgb/astorb.html.

on custom photometry in the sequel to this paper. In this paper, we limit our analysis to cataloged parameters computed by the SDSS photometric pipeline (Lupton et al., 2002).

All objects detected by the SDSS photometric pipeline have two other magnitudes measured (for each of the five bands). First, the PSF magnitude, measured by fitting the point spread function model to the object. For a resolved object such as a comet or galaxy this does not measure all of the flux. Secondly, the Petrosian magnitude is designed to measure extended, resolved objects, and provides a model-independent measurement for resolved galaxies. Petrosian magnitudes are a good choice for nearby bright galaxies, or in our case large comets, but since we will be dealing with comets down to the faint limit where the petrosian magnitudes become less reliable we have made the choice of citing model magnitudes unless otherwise specified. For details on how these magnitudes are measured by SDSS and the advantages of each, please see Stoughton et al. 2002, and references therein.

In addition to assigning magnitudes, whenever the SDSS photometric pipeline encounters a complex object or crowded field it tries to separate that object into component pieces and then reports photometry on these "children" of the original, or "parent," object. This procedure can sometimes take a single resolved object, such as a spiral galaxy, or in our case a comet, and deblend the single object into multiple separate ones, each with its own measured photometry. All the processing steps taken by the photometric pipeline are encoded in an extensive series of processing flags. These flags indicate the status of each object, warn of possible problems with the image itself, and warn of possible problems in the measurement of various quantities associated with

the object [3]. For more details on the technical aspects of SDSS photometry see Stoughton et al. (2002).

2.4 The Initial Sample

Active comets imaged by SDSS are resolved objects, and they are classified by the SDSS photometric pipeline as "galaxy" type objects (Lupton et al., 2001). The first SDSS comet was found serendipitously (Comet C/1999 F2 Dalcanton; Dalcanton et al., 1999), while examining irregular galaxies. Beyond Comet Dalcanton there have been several other serendipitously found comets. Five comets were found by visually inspecting the images as they were processed. This method can only discover the largest objects, because fainter, smaller comets are not usually noticeable in large field images (9×13 arcmin^2). Four comets were also found by examining galaxy spectra. Since active comets are classified as galaxies, SDSS targets them for spectroscopic follow-up. By the time the follow-up spectra were taken (typically within several weeks) the comets were no longer at that position in the sky, resulting in a blank-sky spectrum. Visual inspection of the photometric images from the detection runs revealed that these resolved objects were indeed comets because they exhibited detectable motion during the 5 minutes of imaging.

While these serendipitous detections prove the concept of using SDSS to find and study comets, a robust completeness estimate requires a systematic way to find comets from a single epoch imaging set, using a well-defined single data release of the survey. To do so, we first systematically search for comets

[3] For more details about processing flags see http://www.sdss.org/dr5/products/catalogs/flags.html.

using SDSS galaxy spectroscopy as a second-epoch information source. Due to the design of the SDSS spectroscopic galaxy targeting algorithm (Strauss et al., 2002), this search is flux-limited to $r \leq 17.7$ (petrosian magnitude). Out of the 440,502 ($r \leq 17.7$ petrosian) spectroscopic galaxies in DR5, 6771 have spectra that are inconsistent with galactic spectral classification (representing about 1.5% of the targeted galaxies' spectra). Images of these 6771 objects were visually inspected and 8 comets were found in this sample. We used 50 by 50 arcsecond color composites (jpg) available from the SDSS DR5 site [4] for visual identification. These eight comets contained two new comet detections, as well as all four comets found previously using the same method, and two of the five comets serendipitously identified in images. Comet Dalcanton and another two of the five serendipitous comets that were not found are not in runs that were released as part of DR5. The last comet had been targeted for spectra, but had not yet been followed up at the time of DR5's publication. Table 1 summarizes these results.

While this is a robust method for finding comets in SDSS catalogs, it does not fully utilize SDSS imaging data due to the spectroscopic flux limit. SDSS can be used to reliably find comets to $r < 20$. We utilize the 8 spectroscopically found detections of comets with $r \leq 17.7$ (petrosian) in DR5 as a control group to develop and test a robust method for selecting candidate comets with $r < 20$.

[4] http://cas.sdss.org/astrodr5/en/tools/chart/list.asp

2.5 The Faint Candidates

Five selection criteria based on cataloged parameters measured by the photometric pipeline were developed to significantly reduce the number of objects from the 215 million objects present in DR5. This sample pruning is needed to get a reasonable number of candidates for visual inspection, as was done with the 6771 candidates pre-selected using spectroscopy. This selection algorithm must also preserve the control set of 8 comets which we assume is nearly complete [5] for $r \leq 17.7$ (petrosian). Comets brighter than this are easily recognizable as such, and should only be missed due SDSS image overlap effects (see section 3.1.) Due to the small number of bright comets it is unlikely but possible that a comet was not selected for spectra due to being to close to another object (see Strauss et al. 2002 for details on the limits of completeness for galaxies in the SDSS spectroscopic survey) The selection criteria are:

(1) Objects must be brighter than $r = 20$. This cut is made made to eliminate very faint sources that do not have a sufficient signal-to-noise ratio for reliable visual confirmation of cometary nature.

(2) Objects must be resolved. These are active comets, and thus their coma are imaged as a resolved object, not as a point source. We require that the difference between the psf and model r magnitudes, Δr, be greater than 0.2. This is a slightly less restrictive cut than the one used used for morphological star-galaxy classification by the galaxy spectroscopic targeting algorithm ($\Delta r > 0.3$; Strauss et al., 2002).

[5] Here "nearly complete" means that there are not more than one or two other comets with $r \leq 17.7$ (petrosian) detected in DR5 images, but not recognized as comets.

(3) Objects must have a measured angular velocity greater than $v = 0.04$ deg/day. This limit corresponds to the angular motion of an object in a circular orbit at about 20 AU (roughly the orbit of Uranus). For comparison, main-belt asteroids observed at opposition typically have $v = 0.2$ deg/day (see Figure 14 in Ivezić et al., 2001)

(4) The angular velocity should be significant when compared to its measurement error, and a limit of $v/\sigma_v > 4$ was chosen (see Fig 6 in Ivezić et al., 2001).

(5) All 8 of the control comets have at least one of three `DEBLENDED` flags set by the SDSS photometric pipeline. These flags are set when the pipeline encounters large or complex objects, such as galaxies, crowded fields, galaxy-star blends, and in our case comets. The three flags are: `DEBLENDED_AS_MOVING`, assigned to complex objects that the photometric pipeline identifies as having moved during the 5 minutes of imaging, `DEBLENDED_AT_EDGE`, assigned to larger, complex objects (typically larger than 1 arcmin) that were imaged at the edge of a scan line, and `DEBLENDED_DEGENERATE`, which is assigned to complex objects where the photometric pipeline gave two or more valid deblending solutions. For more details see Section 4.4.3 in Stoughton et al. (2002). The final selection criterion is to require that at least one of these three flags has been set.

In order to check that these selection criteria worked as intended, they were applied to the spectroscopic galaxies ($r \leq 17.7$ petrosian), reducing that sample from 440,502 objects down to just 2262, with no restrictions on spectral classification. Table 2 shows how the application of the five cuts achieve that result. These 2262 images were visually inspected and returned the same 8

comets that had been found through examining the 6771 objects morphologically selected as galaxies, but not matching galaxy-type spectra. No new comets were found in this new set of 2262 candidates, demonstrating that the restriction to inspect only objects that had spectra inconsistent with galaxy classification did not introduce any additional strong incompleteness to the sample (e.g., due to problems with the spectral classification pipeline).

Applying these five cuts to the full DR5 catalog reduces the 215 million object catalog down to 157,996 objects (with $r < 20$). Figure 5 compares the distributions of these 157,996 objects and comets from the control sample in diagrams constructed using the parameters utilized by the selection criteria. Out of these 157,996 objects 106,291 have $17.7 < r < 20$, and may include additional comets that would have been below the magnitude limit for spectroscopic selection.

Objects that are candidate comets must be visually inspected in order to ascertain if they are truly comets. While the initial 5 cuts dramatically reduce the size of DR5 by over a factor of 1000, a sample of $\sim$150,000 objects is still too large to effectively examine by eye. With no second epoch available, tighter constrains must be enforced to further decrease the sample size. Examining the measured properties of the known comets shows several properties that can be taken advantage of to further reduce the sample size. The initial constraints placed on Δr and velocity seem too generous as discussed below. Additionally, the color distribution of the control comets is remarkably narrow (discussed in more detail in sections 2.7 and 3.4). This suggests two possibilities to reduce the false positives: 1) Impose tighter constraints on Δr and minimum velocity of the objects, and 2) further selection criteria can be applied in color-color space based on the narrow range of colors exhibited by the control comets.

2.6 Velocity-Size Selection

Comets in the control sample and all other serendipitously found comets have an angular velocity larger than $v = 0.04$ deg/day. The slowest measured velocity was $v = 0.04$ deg/day, with a median value for the control comets of $v = 0.12$ deg/day. The comets are also significantly more resolved than the lower limit of $\Delta r = 0.2$; the most point-source like has $\Delta r = 1.19$, and the control comets as a group have a median $\Delta r = 1.92$. Therefore, a more restrictive cut on angular velocity and on Δr was made with $v > 0.1$ deg/day and $\Delta r > 1$. Out of the 8 comets from the control sample, 6 pass these more restrictive criteria. The two comets that fail both had measured angular velocities less than 0.1 deg/day (0.04 and 0.09 deg/day respectively). Enforcing these two restrictions reduced the $r < 20$ sample of 159,996 by a factor of four, down to 43,005 objects.

These 43,005 objects were visually inspected and 16 comets were identified by this method: in addition to the 6 comets from the control sample, 10 new comets were found. Two of the 10 new comets also had $r < 17.7$ (petrosian). These two objects were targeted for spectroscopic follow-up, but the actual spectra had not been taken by the time of DR5's release (and thus do not imply an incompleteness of the control sample). The 8 remaining new detections are all fainter ($r > 17.7$ petrosian) than the spectroscopic limit, and could not have been found using the spectroscopic method described in Section 2.4.

2.7 Color Selection

The comets detected by SDSS have a narrow distribution of *g-r, r-i* and *i-z* colors. Table 3 and Figure 6 illustrate this trend, which is discussed in detail in Section 3.4. A three-dimensional color cut was then applied, preserving the majority of known comets, while trying to eliminate as many false positives as possible. We required that $g - r = 0.57 \pm 0.25$, $r - i = 0.23 \pm 0.2$ and $i - z = 0.08 \pm 0.4$. Six out of the 8 control comets satisfy these constraints. One of the control comets that fails these color cuts also failed the size-velocity cuts because of its low angular velocity ($v = 0.04$ deg/day). Figure 7 shows these three cuts in color-color space. Applying these color cuts to the 157,996 candidates left 16,254 color-selected candidates. Those objects were visually inspected, and 14 comets were identified.

When both the selection based on color described here, and the more restrictive velocity and size restrictions (section 2.6) are enforced together on the 157,996 candidate objects ($r < 20$) 4510 objects pass the cuts, containing 13 comets. The one comet selected using the color method that was not also selected by the velocity and size method had a measured angular velocity of $v = 0.09$ deg/day, below the more restrictive limit of 0.1 deg/day.

Utilizing these methods a total of 18 images of 16 active comets were positively identified in DR5. Two comets were imaged twice in the DR5 runs. Including comets that were identified in older runs that have been superseded by the more recent DR5, a total of 22 images of 19 active comets have been found using SDSS.

3 Results

Here we present two initial results based on the sample of 16 active comets (18 detections) found in DR5. SDSS is being treated as a single snap shot of the sky, so to avoid bias we have only used the first observation of the multiply observed comets. The first result is the distribution of active comets with respect to magnitude, and with respect to ecliptic latitude. The second result is the evidence for an extremely narrow distribution of the SDSS colors (*u-g, g-r, r-i, i-z*) of active comets [6] . As mentioned earlier, a comet-by-comet analysis of the physical properties of comets based on custom-made photometric code will be presented in paper II.

3.1 Incompleteness

Using the sample of comets found in DR5 allows us to place lower limits on the distribution of active comets as a function of their magnitude, and position on the sky. The major contribution to the incompleteness of the sample presented here is that the comets must have been "identified" by two different sources. First, the comets must have been properly identified and dealt with in the SDSS software. As mentioned previously, some extended objects may be deblended to the point where the reported properties no longer fall within the bounds of the cuts used in the given selection criteria. This will be shown to be the case for the comets that failed the color cuts in section 3.4.

[6] The RMS of the comet colors are a factor of 3 to 4 times smaller than the color cuts discussed in section 2.7

Second, the comets in DR5 must be identified and confirmed by eye. Two incompleteness issues arise with respect to this. First, the field containing the comet must be present within the DR5 database, as well as the corresponding image stamp. The design of the SDSS imaging survey allows for a small area of overlap between runs adjacent on the sky. If a comet is in one of these "overlap" regions, it is possible for the image of the section of the sky containing the comet to actually belong to another run of a different epoch (when the comet was not at that location). In a sense, for some regions on the sky the SDSS survey contains multi-epoch data, but the image stamps only represent one epoch. This is the reason that comets found in earlier data releases such as C/1999 F2 Dalcanton are no longer present in DR5. The second source of incompleteness because when visually inspecting tens of thousands of images it is possible to miss a few, especially at the fainter magnitudes. As r nears 20 the cometary nature of the comet, easily identifiable at brighter magnitudes becomes much harder to discern.

3.1.1 Tests of Incompleteness

In order to test that this incompleteness is not overwhelming we took a list of currently observable Jupiter family comet orbits maintained by the MPC and calculated, using codes described in Jurić et al. (2002), where their orbits intersected an SDSS field in space and time. These matches were then visually inspected to see if a comet was actually there. Nine comets were matched in DR5 this way. Four of these comets had been found by the selection process detailed in this work. Four more could not have been found by the selection criteria as they were outside the main cuts with all four having $r > 20$. The one remaining comet was missed as it was overlapping with a moderately bright

star, resulting in an object whose velocity and deblending flags excluded it from the main selection criteria. The implied sample completeness based on this test is 80% (4 out of 5 possible detections are recovered by catalog-based search).

3.2 Magnitude Distribution of Active Comets

The sample of comets can be divided into two groups based on our confidence in completeness. The bright comets with $r \leq 18$ represent a much more complete sample than fainter comets. Their transient nature is backed up with a second epoch of observations (the spectroscopic survey), and are bright enough to make visual identification robust.

Making simple linear fits to the cumulative number of comets with respect to a given magnitude allows an estimation both of how many comets could be expected to be found at limits of $r < 20$, such as this study of the SDSS data, and much fainter magnitudes achievable with the next generation of large-scale survey telescopes as discussed in section 4.1. Figure 8 and Table 4 detail these fits to the distribution found in this study. Taking the range of values from the fits we find estimates of between 100 and 260 active comets $r < 20$ on the sky at any time, and between 260 and 2860 active comets with $r < 24$ on the sky at any time. The smaller estimates are the result of fitting only the more incomplete, dim ($r > 18$) sample of comets, while the larger values come from fitting only the brighter ($r < 18$) comets. These estimates are effected not just by this sample's incompleteness, but also the unknown nature of activity levels and timescales for extremely small cometary nuclei, the true size scale for cometary nuclei and the non-uniform distribution of

comets across ecliptic latitude. These factors all should have the tendency to reduce the number of active comets at the faint end.

3.3 Ecliptic Latitude Distribution of Active Comets

Fig 9 illustrates the absolute ecliptic latitude distribution of the 306,751 (9×13 arcminute^2) fields in DR5, the 159,997 candidate comets, and the populations of both the asteroids from MOC3 and the 16 comets found in DR5. Both the asteroid and comet populations are concentrated within low ($\beta < 25$) ecliptic latitudes, as expected for the main belt asteroids and JFCs. Three comets were found at higher ecliptic latitudes, consistent with a second population of higher inclination comets (the LPCs).

These distributions were corrected for the solid angle at a given ecliptic latitude (β) on the sky, as well as the distribution of DR5 observations with respect to β (Figure 10). The cumulative histograms for the asteroid population reaches 90% by $\beta \sim 25$. The JFC component of the comets shows a similar trend, but because the comets also contain a second population (LPCs) there is evidence of a break in the distribution here, as the comet population transitions from one dominated by JFCs at low β to one dominated by LPCs at higher β. All three of the high β comets in DR5 were positively identified as known LPCs by the Minor Planet Center.

3.4 SDSS Colors of Active Comets

As discussed in the context of designing color-based selection cuts in section 2.7, the observed comets show a very narrow color distribution, ($g - r$, $r - i$,

and $i - z$), as reported in Table 3. For example, the root-mean-square scatter for the $g - r$ color is only 0.06 mag, and is only marginally larger than expected measurement errors (about 0.02 mag for point sources, and definitely larger for resolved sources) a 3 to 4 times narrower than the color cuts applied in section 2.7. While four comets significantly deviated (> 0.8 mag) from the median colors ($g - r = 0.57$, $r - i = 0.23$, and $i - z = 0.08$), all four had significant deblending errors. This was typically seen with the coma of the comet being separated into multiple unique objects by the photometric pipeline. The resulting photometry of the comet was thus missing significant flux in at least one filter. Figure 11 illustrates one family of deblended images for a comet that failed the color selection. Even comets that marginally deviated from the median colors showed at least some deblending errors. For example C/1999 F2 Dalcanton deviates from the median color in each of $g - r$, $r - i$, and $i - z$ by about 0.1 mag. In the case of this comet, the coma was properly deblended, but the tail was not.

Despite deblending errors, enough of the comets were treated properly by the pipeline to show that their colors are similar to colors reported by other observers. Transforming the SDSS colors into synthetic $BVRI$, using relations from Ivezić et al. (2007), the colors of the comets in this study fall within the range of colors presented by Lowry et al.. Figure 12 shows the comparison between the comets studies by Lowry et al., and the synthetic $BVRI$ derived for the sample of SDSS comets.

For comparison to other solar system small bodies the SDSS colors of comets are compared to those of the asteroids in the SDSS Moving Object Catalog, in Figure 13. To elucidate a more detailed comparison between the colors of active comets and typical colors for other solar system small bodies, the

reflectance (normalized to the r band) for various sub-populations of the main belt asteroids and the Trojans of Jupiter are shown in Figure 14. We find that the reflectance of active comets most closely matches that of the Jovian Trojans (Szabó et al., 2007). On the onset this result is surprising, as Trojans have solid surfaces, as opposed to a diffuse coma of a comet. This may be indicative of the surface conditions of Trojans, but a detailed analysis of this color similarity is beyond the scope of this work.

3.5 Size-Magnitude Relationship of Active Comets in SDSS

The light from comets has a very different physical origin that the light from galaxies (which are the main sample contaminant), so it might be possible that the surface brightness could serve as an efficient separator of comets and galaxies. We compare the potential surface brightness distribution of confirmed comets and the full candidate sample in Figures 15 and 16. These figures show a comparison of magnitude and the 50% petrosian radius (a measurement of effective angular size made by the SDSS photometric pipeline) for the comets and comet candidates discussed in this paper. Lines of constant surface brightness have been over-plotted for comparison.

As discussed previously the SDSS deblending process lead to some significant photometry errors in several of the comets and may well have had minor effect on even those that fell within the range of median colors. Even without correcting for these deblending issues the comets fit a relationship between magnitude and size. The magnitude of an extended object should be proportional to $-2.5\,log(\Sigma\Theta^2)$, where Σ is a surface brightness, and Θ is the angular size of the comet. Assuming some constant surface brightness, there should

exist a relation of $m = C - 5\,log(\Theta)$. We find that our sample of comets is best fit by a relation of $m = 20.19 - 5\,log(\Theta)$ in the SDSS r band, with Θ expressed in arcseconds.

Hence, it turns out that comets in our sample do have a fairly modest surface brightness distribution, with a root-mean-square scatter around the best-fit m vs. Θ relation of only 0.8 mag in the SDSS r band. However, the value of their median surface brightness distribution is similar to typical surface brightness observed for galaxies, and therefore surface brightness is not a good comet-galaxy separator.

Another potential separator of comets and galaxies is the *shape* of their surface brightness profiles. We quantify this shape using the concentration index (the ratio of 90% and 50% petrosian radii), and show its distribution in Figure 17 for the comets and the candidates. While the distribution of this ratio for confirmed comets is different from the distribution for the whole candidate sample, it is still not as a robust selection parameter compared to the previously discussed cuts on color, size and velocity, and does not even reduce the candidate sample by a factor of 2. It is interesting that the distribution of the petrosian radius ratio for the candidate sample is skewed towards smaller values ($<$1.7), compared to the expected distribution for galaxies, which fall off much more sharply at $<$1.7. As illustrated in Figure 18, one of the reasons why the candidates deviate is that the sample includes a substantial number of saturated stars, in addition to galaxies.

4 Discussion

Most photometric surveys of comets to date have targeted select groups of known comets, and often only in one or two filters. Therefore, the availability of a large multi-filter accurately calibrated and uniformly processed SDSS data set can enable both the discovery of new objects and better characterization of known populations. However, data mining of such a large data set, that includes over hundred million resolved sources, is not trivial. The main difficulty is that over most of the surveyed area SDSS has obtained only single-epoch data, which prevents robust confirmation of motion (with galaxies contributing a large number of false positives).

In this contribution, we used a training sample of comets detected by SDSS that was serendipitously observed, and developed well-controlled selection criteria to greatly reduce the number of plausible comet candidates. After the follow-up visual inspection that rejected false positives, the total sample of SDSS comets presented here includes 19 objects.

In addition to providing information about individual objects, the good understanding of selection effects enables us to study population statistics. We estimated the magnitude distribution down to $r \sim 18$, and found evidence for two populations in the sample ecliptic latitude distribution as can bee seen in the broken slope for the comet population in Figure 10. Perhaps the most surprising result is the extremely narrow color distribution for comets in our sample (e.g. root-mean-square scatter of only $\sim$0.06 mag for the $g - r$ color). While SDSS magnitudes are not designed for comet photometry, no aspect of the processing algorithm should artificially produce such a narrow color dis-

tribution. Nevertheless, in the sequel paper we will present results based on custom photometry optimized for comets. Another interesting result regarding comet colors is their close resemblance of the colors of Jovian Trojans, and their large difference with respect to typical colors of main-belt asteroids. This similarity may hold clues about the possible surface conditions of the Trojans.

4.1 Future Large Scale Surveys

The analysis and conclusions presented here are relevant for upcoming large-scale deep optical surveys such as the Dark Energy Survey (Flaugher & Dark Energy Survey Collaboration, 2007), Pan-STARRS (Kaiser et al., 2002) and the Large Synoptic Survey Telescope (Ivezić et al. 2008, LSST hereafter). The results of section 3.2 can be used to estimate the number of comets that these surveys might discover. For example, assuming LSST's depth for single observation of $r = 24$, we estimate that at any given time LSST should be able to detect about 500 active comets. However, we emphasize that this estimate is uncertain by at least a factor of 2 due to extrapolation of SDSS results to much fainter magnitudes. An additional complication is the unknown nature and duration of the cometary activity for very small bodies as well as the size distribution of nuclei, all problems that these future surveys will help address.

The samples obtained by these upcoming deep multi-epoch surveys will include many observations of the same objects. We used magnitude predictions from the JPL's Horizons database to roughly estimate how long would an object be observable (see Figure 19). If reliable, these magnitude predictions imply that some fraction of comets would be bright enough to be detectable by LSST at any point in their orbits. Assuming conservatively that a comet would

be observable only for about 6 months (and ignoring the fact that periodic comets would sometimes "come back"), we estimate that the LSST sample collected during a 10-year long survey would include about 10,000 objects. On average, these objects would be observed about 50 times, but for some of them the LSST data set could include as many as 1000 detections. This volume of observations will allow for the detailed study of time evolution of cometary activity while active and establish light curves for the nuclei while inactive. Another advantage of these upcoming surveys that will obtain repeated observations is the opportunity to utilize image differencing techniques. Not only would this technique enable robust searches for moving resolved sources, but would also help with detailed measurement of their profiles without the contamination by background stars and galaxies.

Acknowledgments

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS Web Site is http://www.sdss.org/.

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study,

the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington. .

References

Abazajian, K. et al. 2004, Astrophys. J., 128, 502

—. 2005, Astrophys. J., 129, 1755

—. 2003, Astrophys. J., 126, 2081

Adelman-McCarthy, J. K. et al. 2007, ApJS, 172, 634

—. 2006, ApJS, 162, 38

A'Hearn, M. F., Millis, R. L., Schleicher, D. G., Osip, D. J., & Birch, P. V. 1995, Icarus, 118, 223

Dalcanton, J., Kent, S., Okamura, S., Williams, G. V., Tichy, M., Moravec, Z., Koff, R. A., & Magnier, G. 1999, IAU Circ., 7194, 1

Flaugher, B., & Dark Energy Survey Collaboration. 2007, BASS, 209, 22.01

Fukugita, M., Ichikawa, T., Gunn, J. E., Doi, M., Shimasaku, K., & Schneider, D. P. 1996, Astrophys. J., 111, 1748

Gunn, J. E. et al. 1998, Astrophys. J., 116, 3040

—. 2006, Astrophys. J., 131, 2332

Hogg, D. W., Finkbeiner, D. P., Schlegel, D. J., & Gunn, J. E. 2001, Astrophys. J., 122, 2129

Hsieh, H. H., & Jewitt, D. 2006, Science, 312, 561

Ivezić, Ž. et al. 2003, Memorie della Societa Astronomica Italiana, 74, 978

—. 2004, Astronomische Nachrichten, 325, 583

—. 2007, Astrophys. J., 134, 973

—. 2001, Astrophys. J., 122, 2749

Ivezić, Ž., Tyson, J. A., Allsman, R., Andrew, J., Angel, R., & for the LSST Collaboration. 2008, ArXiv e-prints, 0805.2366

Jurić, M. et al. 2002, Astrophys. J., 124, 1776

Kaiser, N. et al. 2002, in Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, Vol. 4836, Survey and Other Telescope Technologies and Discoveries. Edited by Tyson, J. Anthony; Wolff, Sidney. Proceedings of the SPIE, Volume 4836, pp. 154-164 (2002)., ed. J. A. Tyson & S. Wolff, 154–164

Levison, H. F., & Duncan, M. J. 1997, Icarus, 127, 13

Lowry, S. C., & Fitzsimmons, A. 2001, A&A, 365, 204

Lowry, S. C., Fitzsimmons, A., Cartwright, I. M., & Williams, I. P. 1999, A&A, 349, 649

Lowry, S. C., Fitzsimmons, A., & Collander-Brown, S. 2003, A&A, 397, 329

Lupton, R., Gunn, J. E., Ivezić, Ž., Knapp, G. R., & Kent, S. 2001, in Astronomical Society of the Pacific Conference Series, Vol. 238, Astronomical Data Analysis Software and Systems X, ed. F. R. Harnden, Jr., F. A. Primini, & H. E. Payne, 269–+

Lupton, R. H., Ivezić, Ž., Gunn, J. E., Knapp, G., Strauss, M. A., & Yasuda, N. 2002, in Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, Vol. 4836, Survey and Other Telescope Technologies and Discoveries. Edited by Tyson, J. Anthony; Wolff, Sidney. Proceedings of the SPIE, Volume 4836, pp. 350-356 (2002)., ed. J. A. Tyson

& S. Wolff, 350–356

Pier, J. R., Munn, J. A., Hindsley, R. B., Hennessy, G. S., Kent, S. M., Lupton, R. H., & Ivezić, Ž. 2003, Astrophys. J., 125, 1559

Scranton, R. et al. 2002, ApJ, 579, 48

Smith, J. A. et al. 2002, Astrophys. J., 123, 2121

Stokes, G. H., Evans, J. B., Viggh, H. E. M., Shelly, F. C., & Pearce, E. C. 2000, Icarus, 148, 21

Stoughton, C. et al. 2002, Astrophys. J., 123, 485

Strauss, M. A. et al. 2002, Astrophys. J., 124, 1810

Szabó, G. M., Ivezić, Ž., Jurić, M., & Lupton, R. 2007, MNRAS, 377, 1393

Tucker, D. L. et al. 2006, Astronomische Nachrichten, 327, 821

York, D. G. et al. 2000, Astrophys. J., 120, 1579

Method	Found	In Spectra	Why missed by spectra?
Dalcanton	1	0	Not in DR5
Visual	5	2	1 future target, 2 not in DR5
Accidental Spectra	4	4	all found
Spectra check	8	8	self reference - 2 new ones
Total	12	8	1 target and 3 not in DR5

Table 1
The comets initially found in the SDSS along with their method of discovery. Of these 12 detections, 4 do not have spectroscopic follow-up.

Criteria	Cut	DR5 Objects	DR5 Spectroscopic Galaxies
DR5	None	approx 215 million	440,502
Bright	$r < 20$	69.9 million	440,502
Resolved	$(psf - model)r > 0.2$	18.6 million	440,502
Moving	$v > 0.04\ degrees/day$	1.6 million	6995
Significant	$v/\sigma_v > 4$	1 million	4244
Flag Set	at least one flag set	157,996	2262

Table 2
The initial selection criteria applied to DR5 and the subset of spectroscopic galaxies. Later criteria are cumulative with prior ones.

Comet	r	$u-g$	$g-r$	$r-i$	$i-z$	°/day	Δr	Δ	R
30P Reinmuth	14.85	1.60	0.57	0.32	0.19	0.42	2.2	1.46	1.88
46P/Wirtiran †	18.84	−1.25	3.39	0.22	0.12	0.32	1.63	1.64	2.60
50P/Arend	18.13	1.72	0.57	0.08	0.27	0.26	2.35	2.57	2.81
62P/Tsuchinshan	15.04	1.27	0.45	0.18	0.07	0.18	3.91	0.95	1.90
65P/Gunn (first)	17.13	1.68	0.58	0.21	0.08	0.12	1.75	3.55	4.34
65P/Gunn (second)	17.15	1.57	0.55	0.21	0.09	0.11	1.9	3.56	4.34
67P/Churyumov-Gerasimenko	14.29	1.83	0.68	0.24	0.06	0.29	3.15	1.59	1.84
69P/Taylor	15.59	1.51	0.60	0.25	0.06	0.29	3.19	1.13	1.95
70P/Kojima	16.68	1.47	0.64	0.23	0.21	0.17	1.97	1.63	2.59
C/1999 F2 Dalcanton	15.18	1.76	0.62	0.31	−0.18	0.06	2.29	4.34	5.00
C/2000 K2	16.82	1.55	0.55	0.22	0.10	0.09	1.74	6.00	5.25
C/2000 QJ46 (first)	17.44	1.55	0.57	0.23	0.12	0.16	1.64	1.16	2.17
C/2000 QJ46 (second)	19.39	1.04	0.56	0.41	0.09	0.18	1.47	3.32	3.66
C/2000 SV74	14.68	1.66	0.52	0.21	0.04	0.11	3.55	4.02	4.38
C/2000 Y2 Skiff	16.71	1.58	0.55	0.18	0.09	0.12	1.19	1.82	2.79
C/2001 RX14 †	12.62	1.24	1.43	0.10	−0.33	0.04	2.76	1.62	2.10
C/2002 O7 (first)	19.37	1.35	0.58	0.14	−0.21	0.20	1.23	5.65	5.89
C/2002 O7 (second) †	15.24	0.60	2.07	−0.27	−0.71	0.15	2.94	2.50	3.04
C/2002 T5	18.33	1.52	0.66	0.23	0.09	0.12	1.97	4.35	5.00
P/1999 V1 Catalina	17.36	1.73	0.58	0.23	0.09	0.13	1.87	2.16	3.10
Unk1	18.20	2.71	0.48	0.36	0.24	0.26	1.97	–	–
Unk2 †	18.75	1.86	0.98	0.14	−1.49	0.27	2.30	–	–
Median Color		1.62	0.57	0.24	0.08				
Std Dev		0.34	0.06	0.08	0.12				

Table 3

The properties of active comets identified in SDSS. The table lists r band magnitude, SDSS colors, measured velocity, the difference between the model and psf magnitudes (Δr). The geocentric distance (Δ) and the heliocentric distance (R) in AU at the time of observation are also cited, except for the two comets designated by *Unk1* and *Unk2* which were not matched to any known object (comets or asteroids). The designations (first) and (second) refer to the chronological order of comets with multiple detections. Comets marked by a † suffer from significant deblending issues, and were not included in the calculation of median and standard deviation for the sample color distributions.

Fit	$r < 20$ 1/5 sky	$r < 20$ all sky	$r < 24$ 1/5 sky	$r < 24$ all sky
bright: $r < 18$	52	260	572	2860
all: $r < 20$	28	140	218	1090
dim: $r > 18$ and $r < 20$	20	100	52	260

Table 4
For the three fits in Fig 8 the total number of comets expected at both 20th and 24th magnitudes, for a 1/5 sky survey (like SDSS DR5) and on the whole sky.

filter	fit (C)
u	22.64 ± 1.43
g	21.17 ± 0.88
r	20.19 ± 0.74
i	20.04 ± 0.73
z	19.97 ± 0.98

Table 5
Fits of the comets to $m = C - 5log(\Theta)$, as seen in Fig. 16 and Fig. 15 with standard deviation of the data around the median.

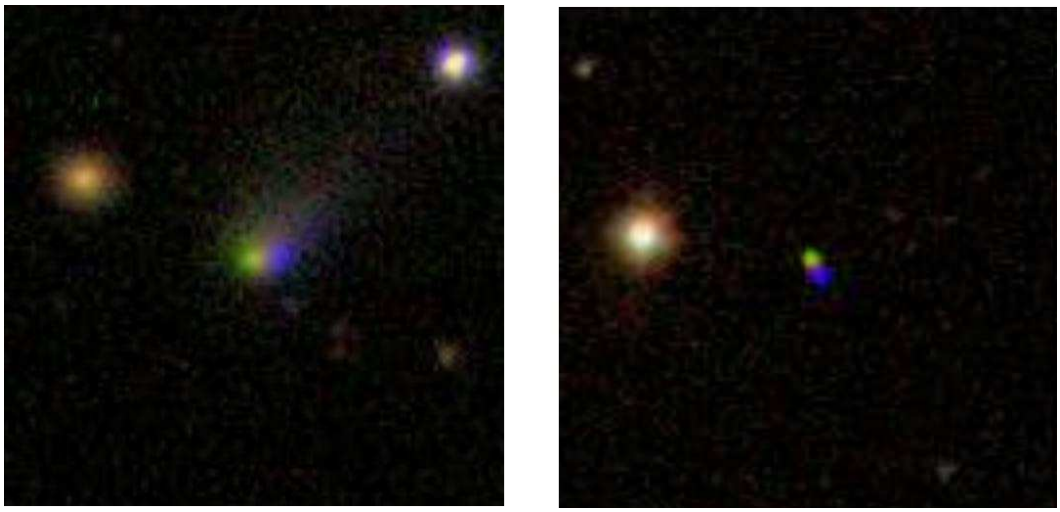

Fig. 1. A three color (g,r, and i) composite of comet P/1999 V1 Catalina (left) and asteroid 2006 SP_{363} (right) as imaged by SDSS. Each image is a 50" × 50" square.

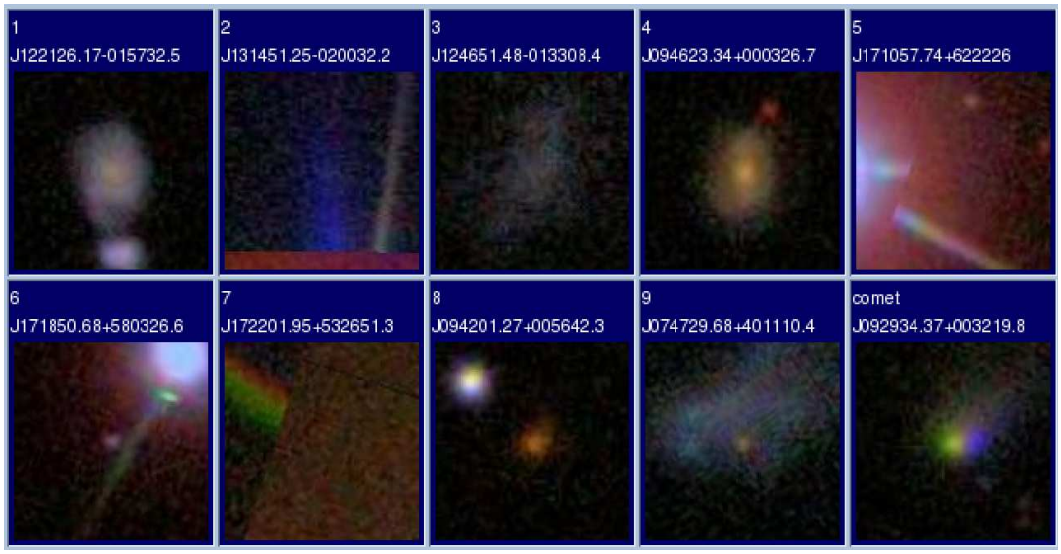

Fig. 2. An image collage comparing the appearance of comet P/1999 V1 Catalina (lower right) with a sample of false positives in the SDSS skyserver. All image stamps are 24" × 24". The majority of the false positives come from galaxies, but imaging issues (in the form of diffraction spikes and "ghosting" from a bright source out of the frame) also contribute.

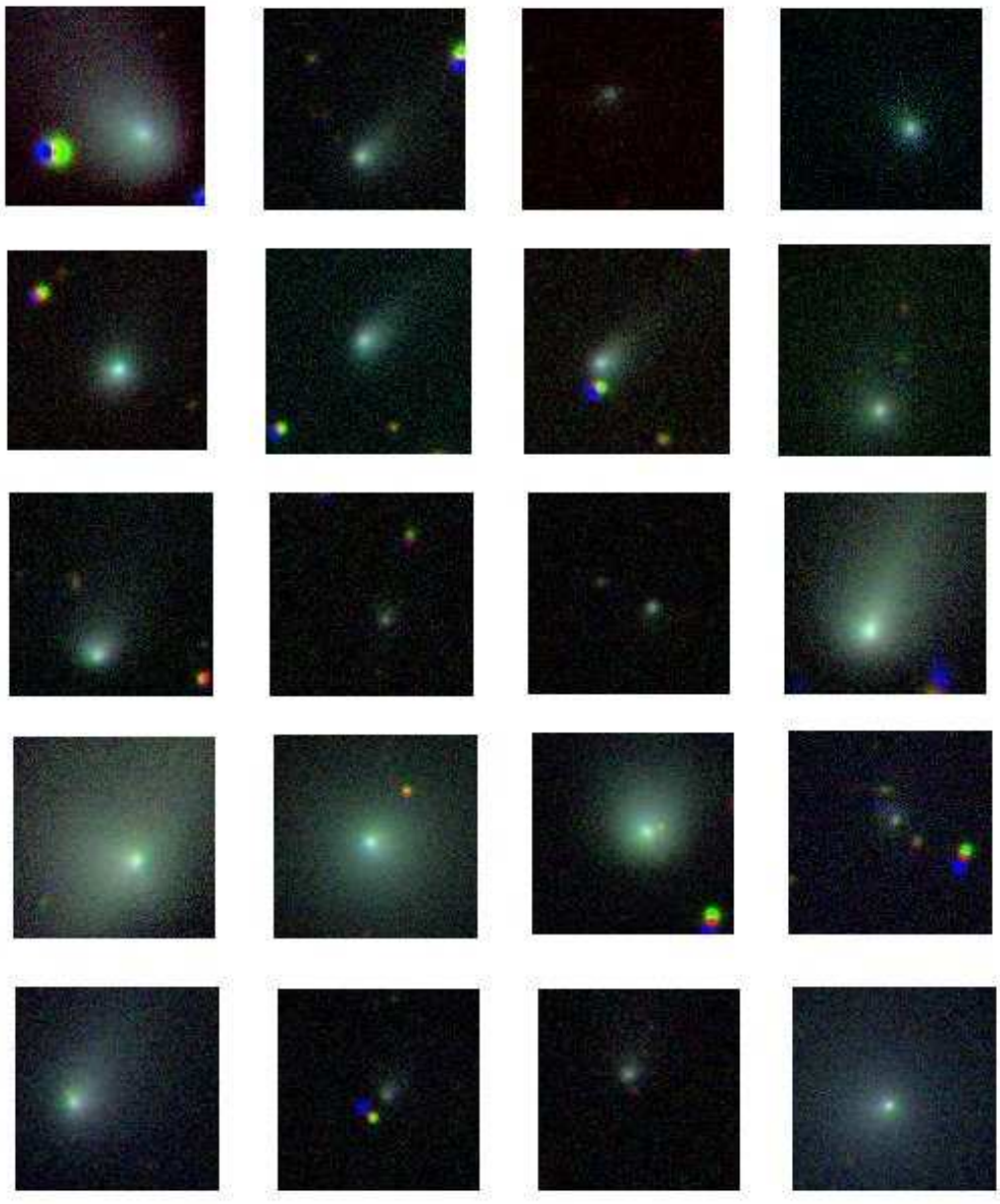

Fig. 3. An image collage of 20 images of 17 comets discussed in the paper. The color images are RGB color composites of the i, r and g filters of the comets. All image stamps are approximately 40" $\times$ 40". These RGB images were made by registering the images to the motion of the comets, rather than the position of the background stars (as was done for Fig. 1).

Fig. 4. An RGB color composite of the i, r and g filters of the large comet C/2001 RX14 (LINEAR). This is an entire SDSS field and is approximately 811" $\times$ 590".

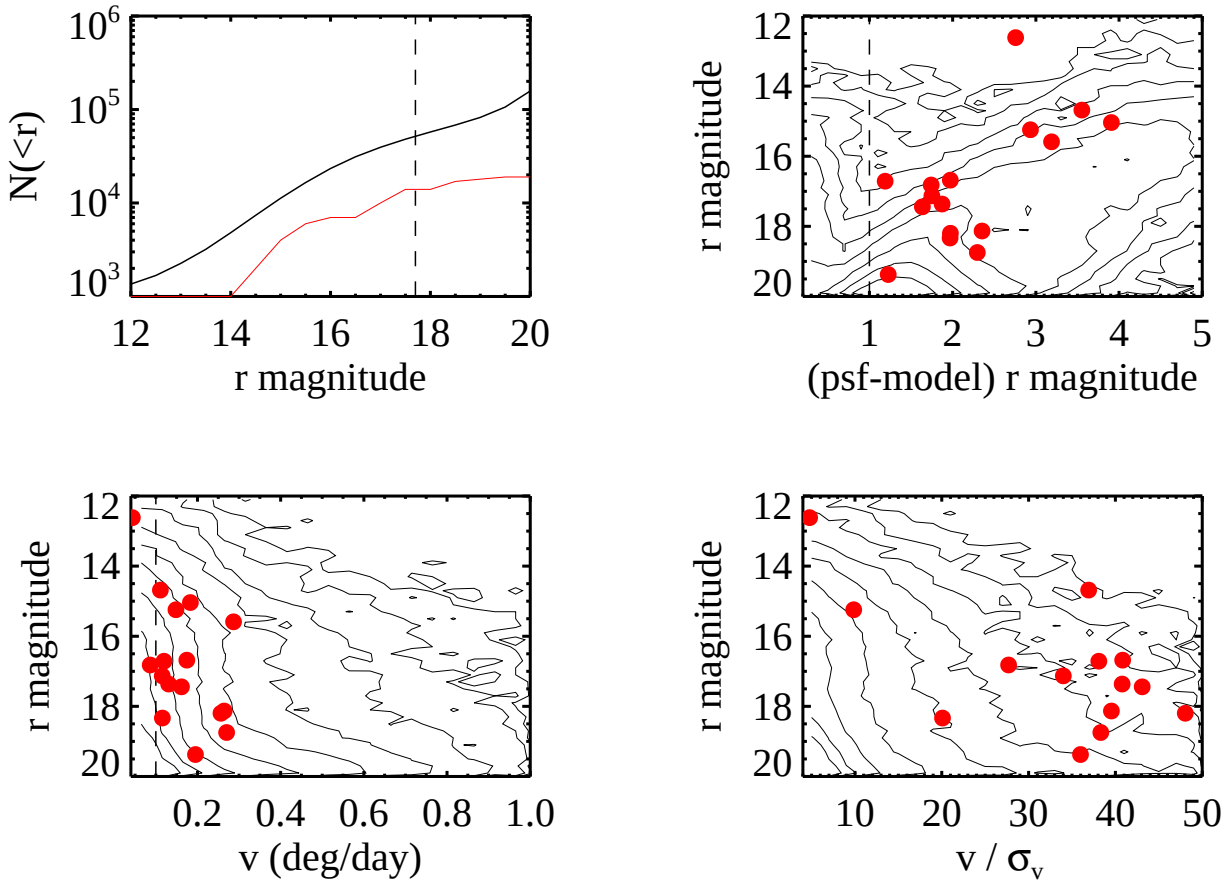

Fig. 5. Top left: Cumulative r-band magnitude distribution of the 157,996 candidate objects (thick line), and the SDSS comets (thin line),the dashed line at r = 17.77 shows the limit of SDSS spectroscopy; the other three plots show the distribution of measured properties used in the selection criteria for the 157,996 candidate objects (contours) and the control comets from spectroscopy (symbols). The vertical dashed lines in the top right and bottom left panels show the more restrictive size and velocity cuts discussed in section 2.6.

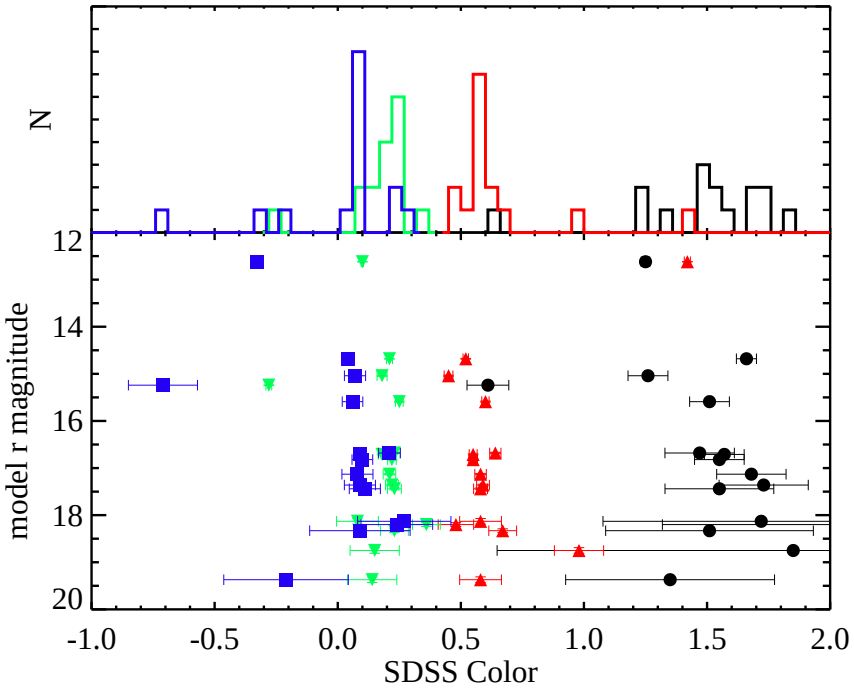

Fig. 6. The Color-magnitude distribution of the DR5 comets with 2-σ error bars as determined by the photometric pipeline. The plotted symbols are: $\bigcirc$ $u-g$ (black), $\triangle$ $g-r$ (red), $\triangledown$ $r-i$ (green), and $\square$ $i-z$ (blue). Plotted above is the histogram of the color distribution. The large outliers are comets with significant deblending errors as discussed in section 3.4

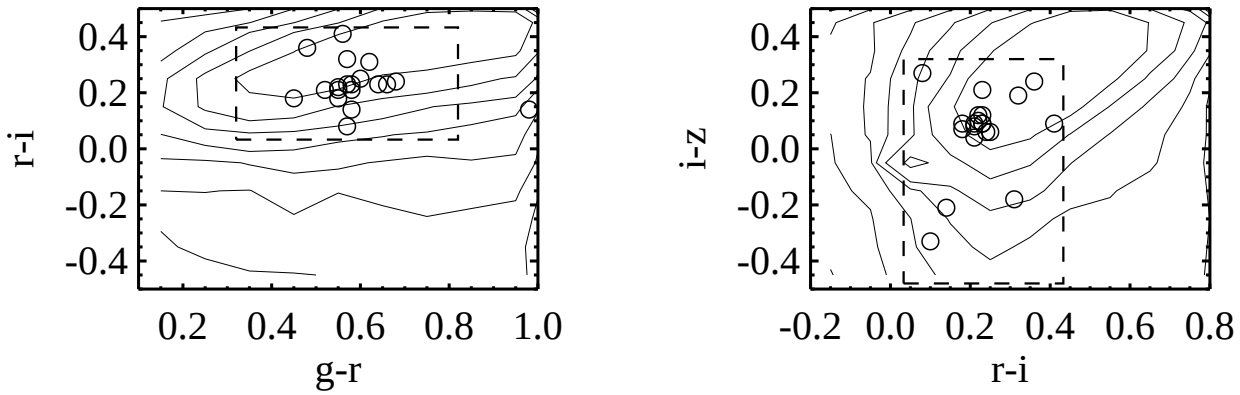

Fig. 7. The dashed-line rectangles show the cuts made in color-color space, shown on top of contours of the 157,996 candidates. Comets are shown by open circles

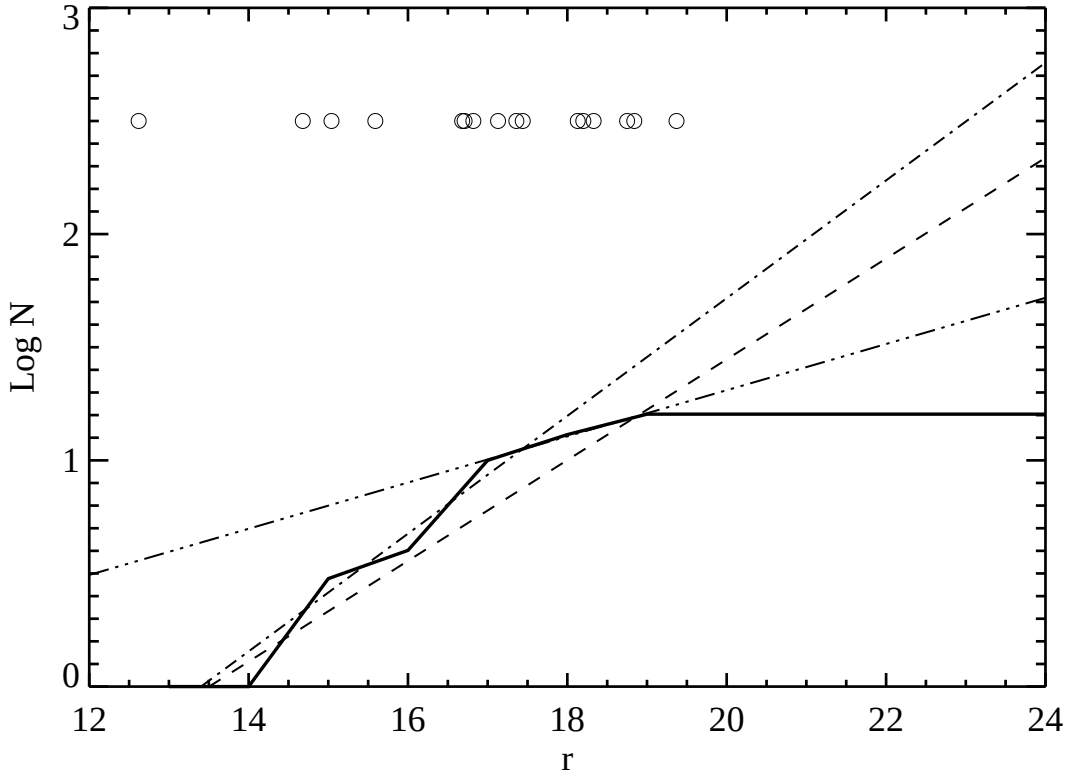

Fig. 8. Log of the cumulative number of comets found in DR5 with respect to r magnitude (solid line). Three other lines represent fits to (dashed) the total ($r < 20$) magnitude range, (dash-dot) the bright ($r < 18$) range, and (dash-triple dot) the faint ($r > 18$) range. The actual comet magnitudes are represented by open circles above the plot.

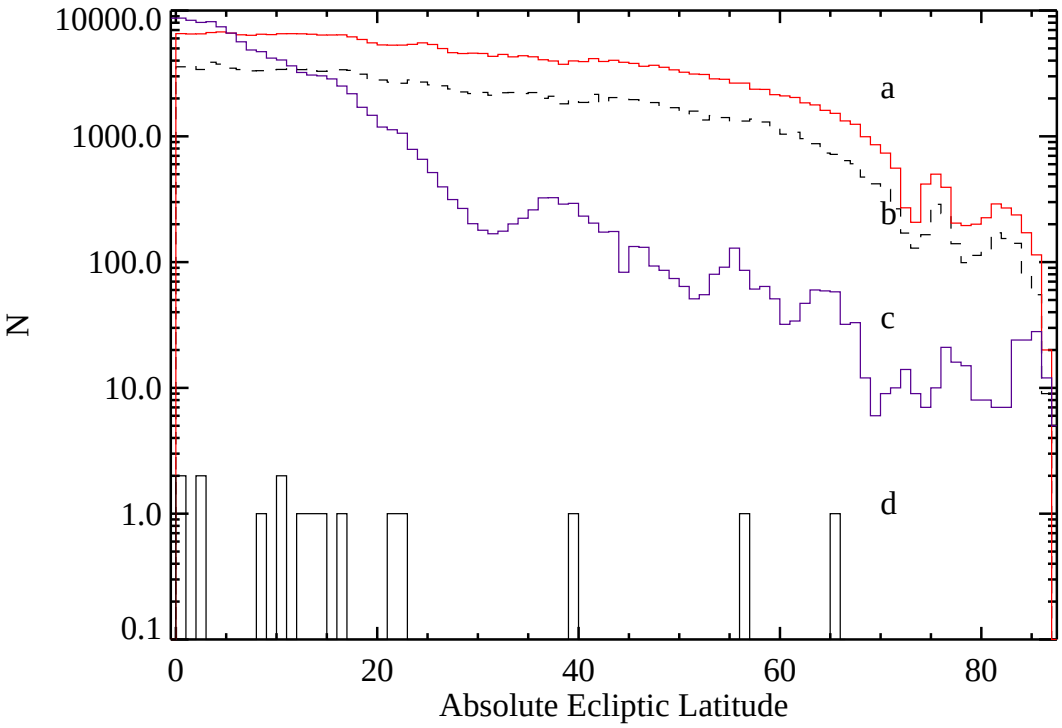

Fig. 9. The distribution of (a) all DR5 fields, (b) the 157,996 candidate comets, (c) the asteroids in MOC3, and (d) the 16 DR5 comets with respect to absolute ecliptic latitude. Two populations of comets can be seen. All known Jupiter family comets in the sample can be found at low β, and the three comets observed at high β were identified as long period comets by the Minor Planet Center.

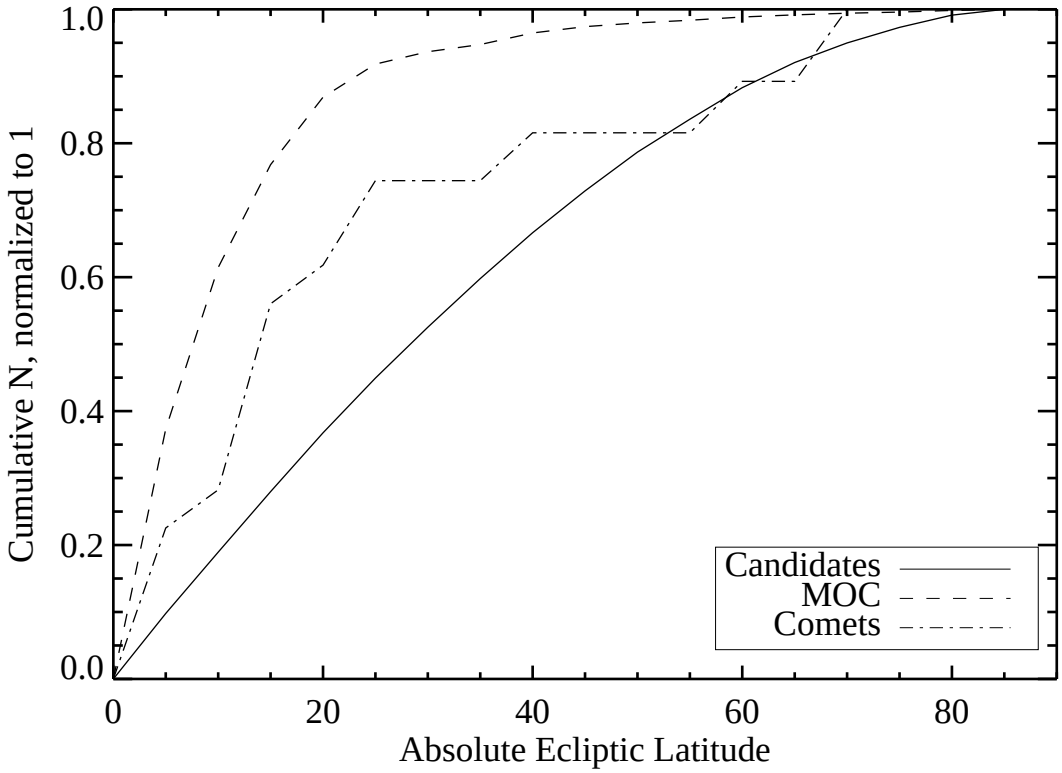

Fig. 10. The absolute ecliptic latitude distributions from Fig 9, corrected for the latitude distribution of DR5 fields, and the differential solid angle at a given ecliptic latitude. Unlike the asteroids in the MOC the comet population demonstrates a broken slope at around $\beta = 25$ showing evidence for two comet populations.

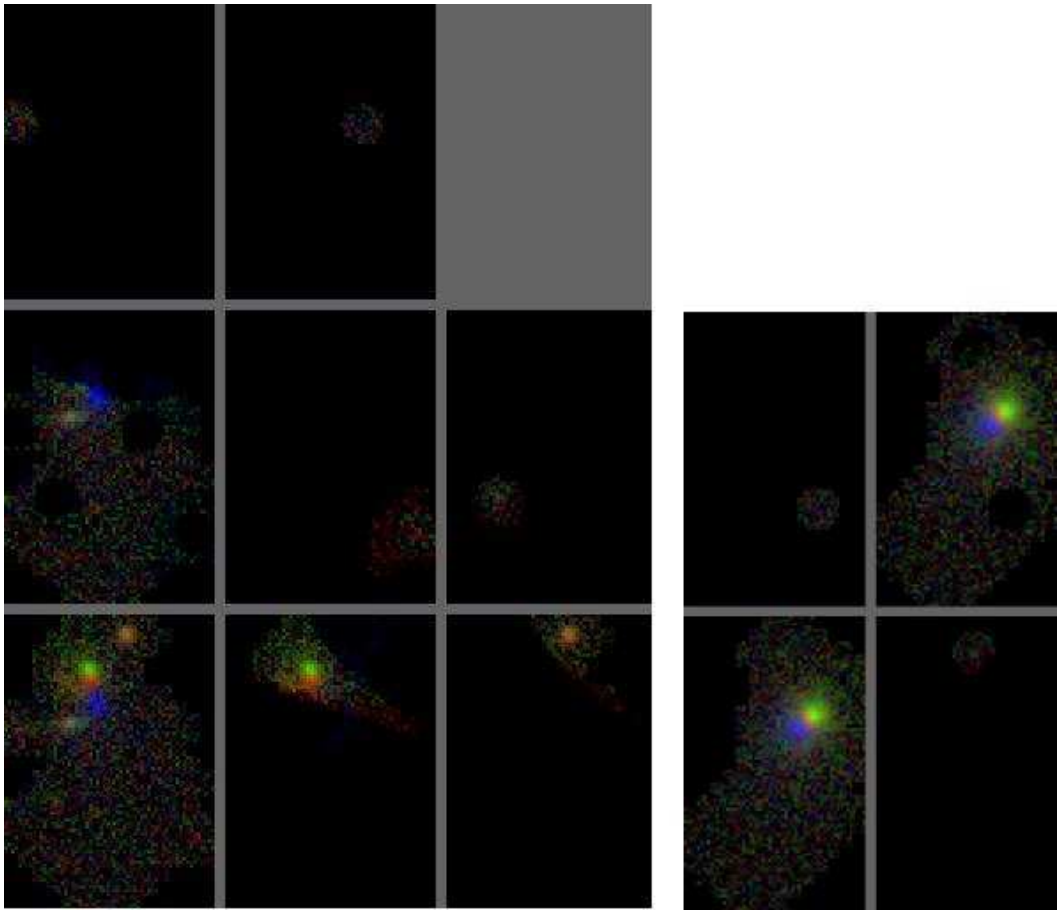

Fig. 11. Left: An example of the badly deblended comet 46P/Wirtiran. The colors represent the g, r and i filters. The lower left image is the composite (gri) image, and the others are the separately treated objects that the deblender has divided the original image into. Notice that the head of the comet has been deblended into two separate objects (middle row left and bottom row center). Right: A similar figure showing the first observation of comet C/2000 QJ46 (LINEAR) which was not pulled apart in the deblending process.

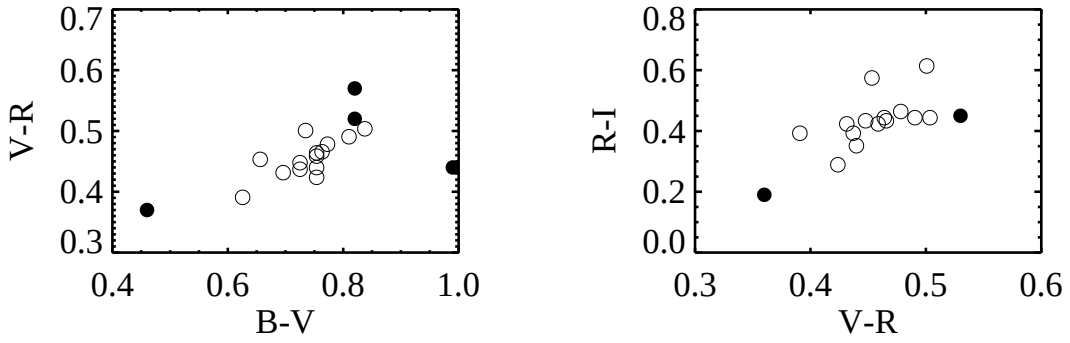

Fig. 12. The colors of the SDSS comets transformed into *BVRI* (open circles) and compared to the colors of active comets cited in Lowry et al. (1999), Lowry & Fitzsimmons (2001), and Lowry et al. (2003) (filled circles). The transformation from SDSS colors to *BVRI* is given in Ivezić et al. (2007)

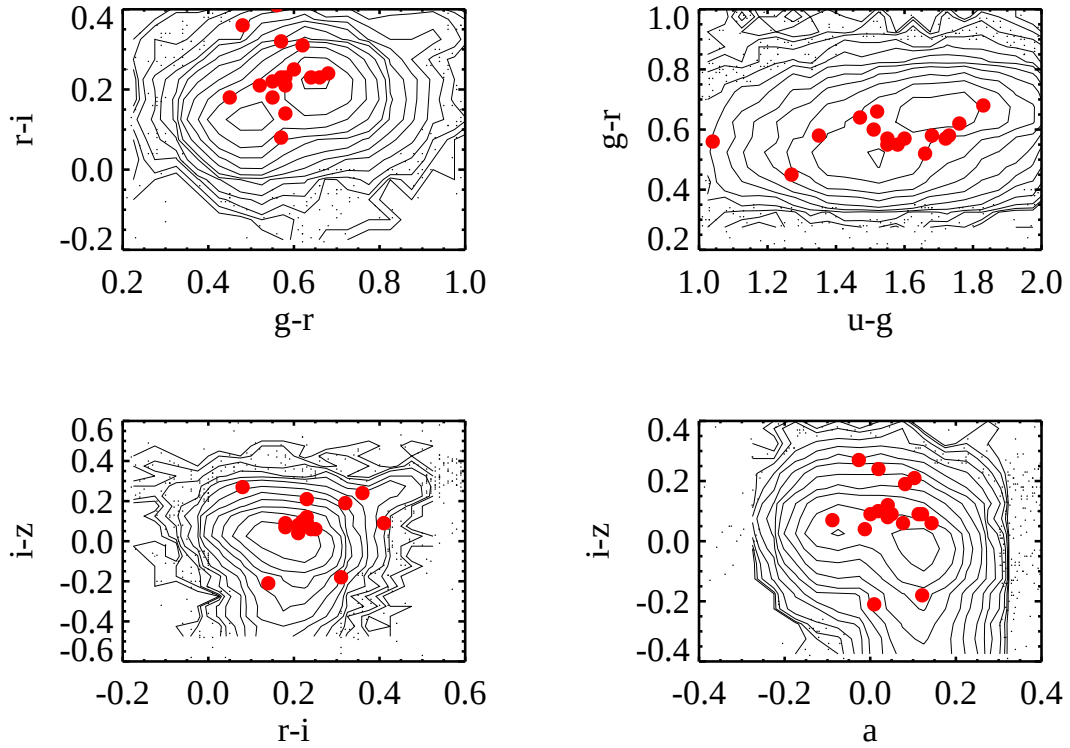

Fig. 13. The color distribution of comets (filled circles) vs. the colors of the asteroids listed in the MOC. "a" is the principle component color in the MOC defined as $a = 0.89(g-r) + 0.45(r-i) - 0.57$ (see Ivezić et al. 2001)

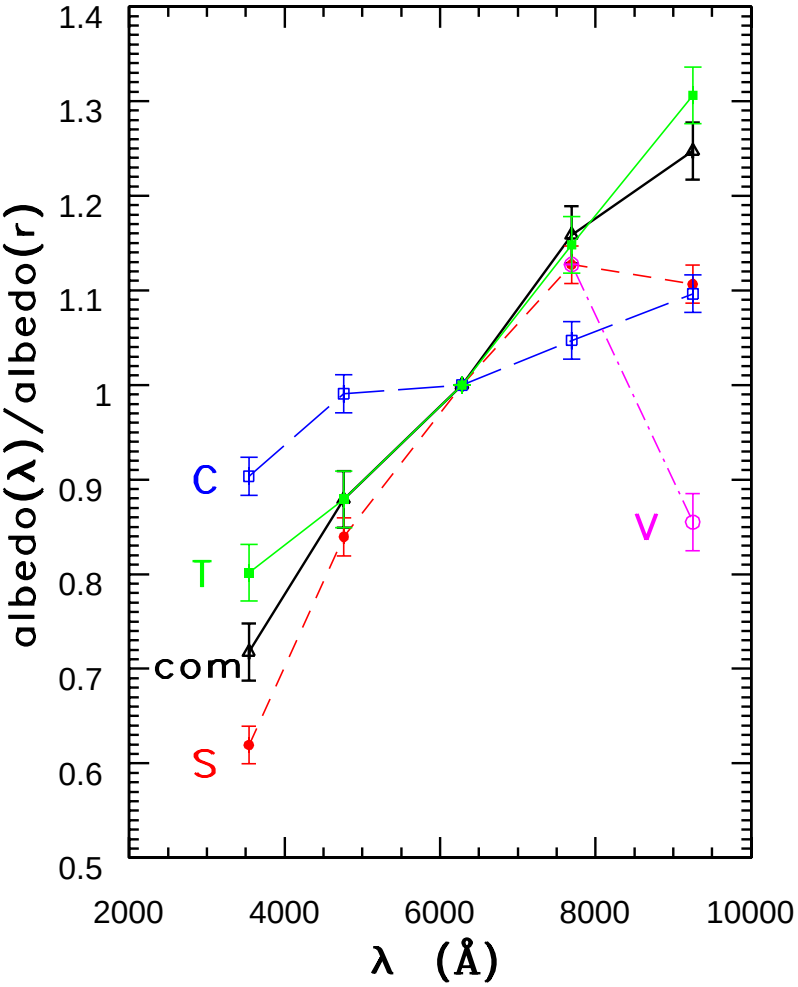

Fig. 14. A comparison of the relative albedo for comets (black triangles) and the relative albedo for Jovian Trojans (green filled squares) and the three dominant main-belt color types (C type: open blue squares; S type: red dots; V type: magenta open circle; the V type differs from the S type only at the longest displayed wavelength). Due to large sample sizes, errors reflect systematic absolute uncertainties in SDSS photometric calibration. Because the same systematic uncertainty applies to all curves, their differences are highly significant (the relative positions of data points are accurate to about 0.01). The albedo curve for comets is more similar to the albedo curve of Jovian Trojans than to albedo curves for main-belt asteroids.

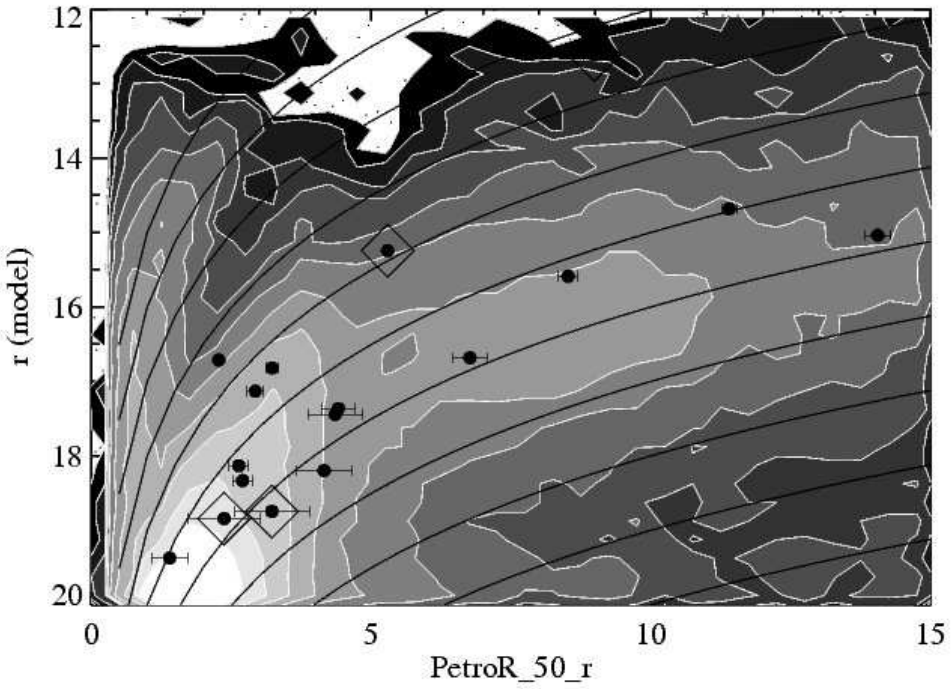

Fig. 15. The relationship between the magnitude of the observed comets (filled circles, with those suffering from major deblending errors surrounded by open diamonds) and their angular radius, Θ, taken as their measured 50% petrosian radius (with 2σ error bars) in the r band. The lines represent lines of constant surface brightness, $r = C - 5log(\Theta)$ for constant values of C ranging from 15 to 25. The best fit for the comets is $C = 20.19 \pm 0.74$ (see Table 5).

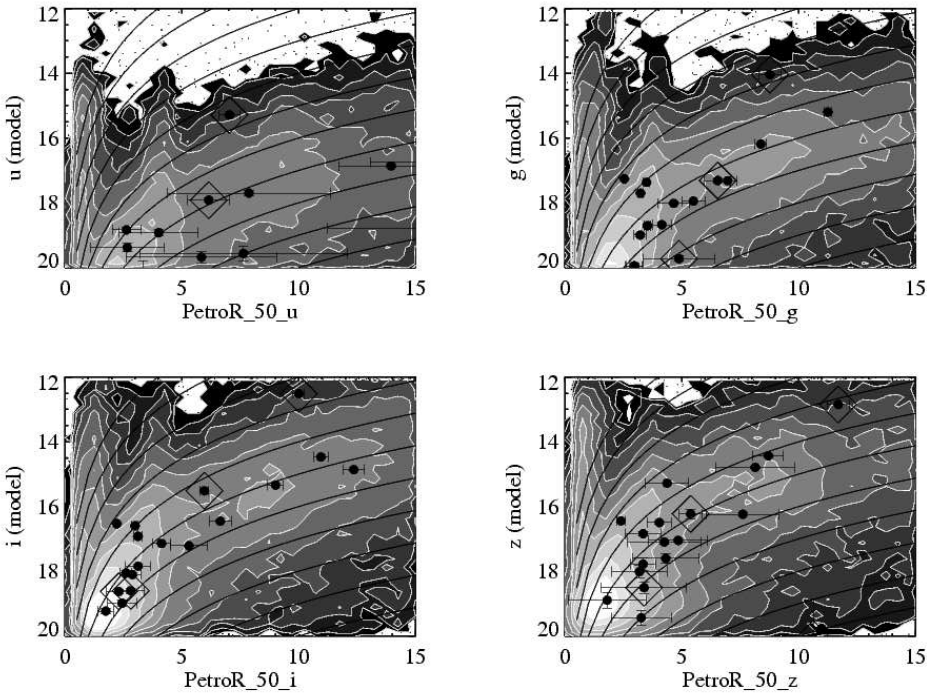

Fig. 16. Analogous to Fig 15, except for u,g,r and z bands. Table 5 summarizes the fits for all five bands.

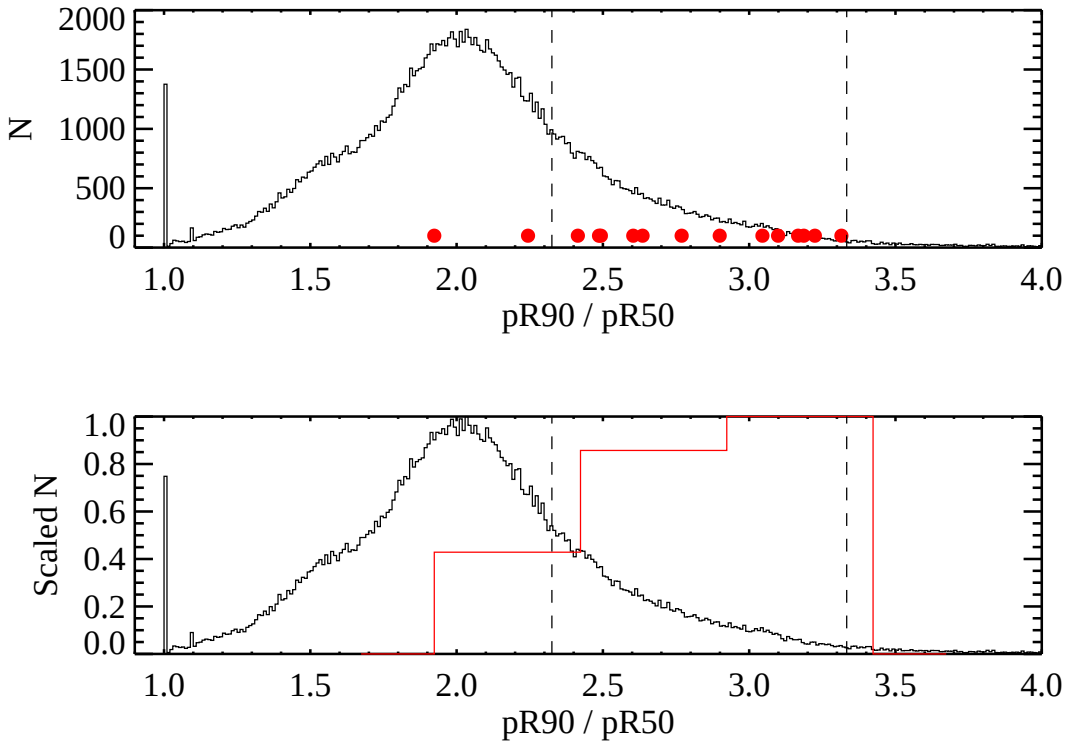

Fig. 17. Histogram of the ratio of the 90% to 50% petrosian radius (pR90 and pR50) measured for each of the 157,996 DR5 candidates. The DR5 comets are shown as red dots. Two dashed vertical lines represent what would be expected from the two model fits (exponential and de Vaucouileurs). The singular spike at pR90/pR50 = 1 is from diffraction spikes.

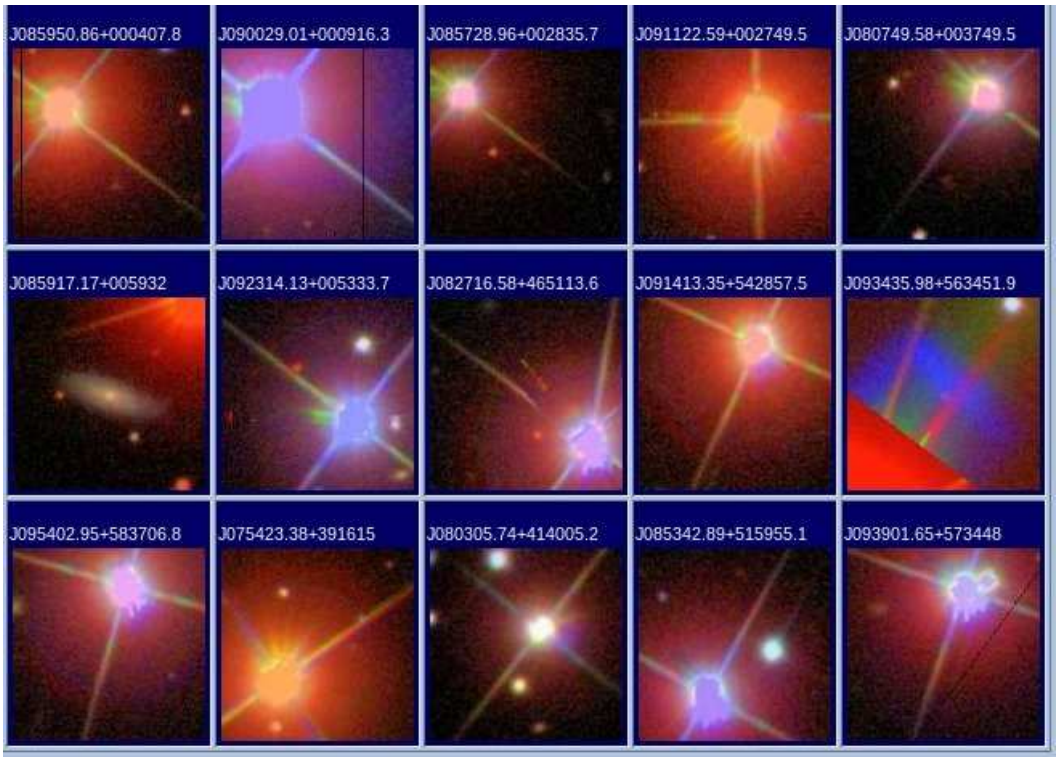

Fig. 18. Examples of candidate objects with $8 <$ PetroR50 < 10 and $16 < r < 17$. (See Fig 15). These don't look significantly different than those objects that populate the feature with $1 <$ PetroR50 < 2, and $15 < r < 16$.

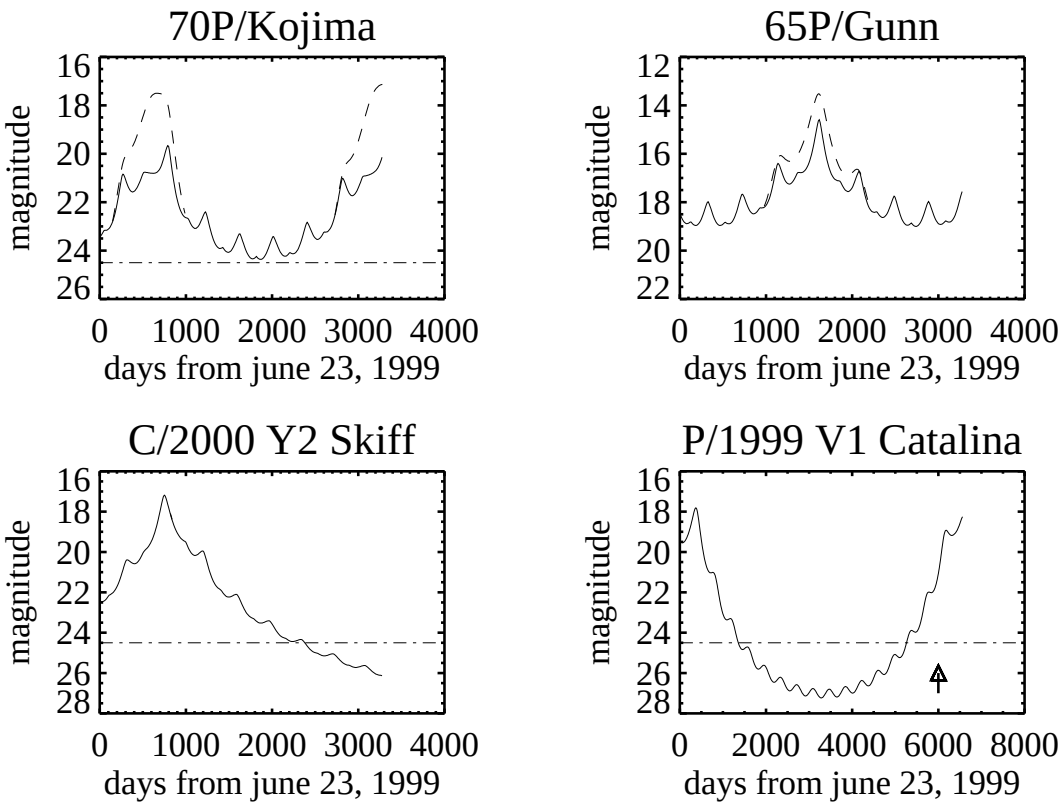

Fig. 19. The predicted visual magnitude of 4 comets from this study over at least a portion of their orbits. The comet will be inactive (asteroidal in appearance)for most of its orbit, and the solid lines represent the predicted nuclear magnitude of the comet. The rising dashed lines show the predicted magnitude the comet will reach while active during that part of it orbit. All magnitudes were predicted by the Jet Propulsion Laboratory's Horizons database (http://ssd.jpl.nasa.gov/?horizons). The horizontal dash-dot line at 24.5th magnitude represents the limiting magnitude for a single observation with LSST, which should see first light at around day 6000 (2015). As can be seen for both comets Kojima and Gunn, LSST will be able to detect these comets at all observable points in their orbits, and even for very long period comets like Skiff LSST will be able to greatly enhance the period of observable time for these objects.