SLAC-TN-04-072
September 2004

A Search for Variable Sources in the Sloan Digital Sky Survey

Allison Widhalm[a] & Masao Sako[b]

[a]Department of Physics and Astronomy,University of Southern California, Los Angeles, CA 90089-0484
[b]KIPAC/SLAC, M/S 29, Sand Hill Rd., Menlo Park, CA 94025

The Sloan Digital Sky Survey (SDSS) is an ambitious on-going project, which is mapping in detail one quarter of the entire sky to study the properties of over a 100 million astronomical objects. A fraction of the sky is observed more than once, which allows us to search for transient objects (sources that vary in brightness), some of which may be associated with supernovae, afterglows of gamma-ray bursts, as well as activity of material near a super-massive black hole in a quasar. The project will involve photometric analyses of a large amount of data from the SDSS archive. Accurate photometric measurements over the entire survey area will allow us to search for transient events that occur on timescales ranging from days to years. We will identify the types of objects, and possibly execute follow-up observations with other observatories.

1. Introduction

Understanding the properties of the optical sky in the temporal domain is crucial for planning and optimizing observation strategies for future projects, such as the Large Synoptic Survey Telescope (LSST) and the Supernova Acceleration Probe (SNAP), which will both survey large areas of the sky in the optical and infrared band and on various timescales. At this point in history, relatively little is known about optically faint variable sources. This preliminary paper enforces this idea, with evidence from the Sloan Digital Sky Survey, that further research is worth while.

The Sloan Digital Sky Survey (SDSS) covers ∼10000 square degrees of the sky in 5 photometric bands (u, g, r, i, and z) that span a wavelength interval of 3000 - 9500 Angstroms up to a depth of r ∼ 22 magnitude. Observations are performed in stripes, which have substantial spatial overlap. An estimated 40% of the northern survey area is observed twice (Ivezic et al. 2003), which allows us to search for variable sources on various timescales. Although photometric data on any given object is available only during two epochs, the large area covered by the survey allows us to statistically characterize the nature of their variability. Currently the SDSS is the largest survey that will allow us to scan the sky up to a depth of r ∼ 22. It serves as a mini-survey of future projects such as the LSST and SNAP.

In this paper, we discuss our initial attempts to characterize the properties of variable sources in the overlap sky regions of the SDSS data. We perform photometry on sources that lie in these regions, identify the sources using their color information, and search for variability. While time-dilated variability in point sources is what we are searching for, variability between the five photometric bands provides key insight into the characterization of sources. Most sources are expected to be quasars and stars, but other types of sources are expected as well. These other sources may prove important in followup studies, particularly with new devices such as the LSST and SNAP.

2. SDSS Observations and Analyses

The Sloan Digital Sky Survey (SDSS) covers the sky in strips that consist of 6 columns of CCDs (camcol). The "camcols" are separated by gaps that are filled in by an offset neighboring strip. Two strips cover a stripe. We have examined the sky overlap regions of Stripe 34 Camcol 6 and Stripe 35 Camcol 1 of the SDSS Data Release 2 (DR2). These consist of a total of 21 runs and 3180 fields. A run consists of a continuous scan along a stripe and is divided into fields of equal sky area. Each field is approximately 13.5' × 8.9' in size and photometric data in five filters (u, g, r, i, and z) are available. The 95% point source detection repeatability magnitude limits in each of the filters are u=22.0, g=22.2, r=22.2, i=21.3,

Work supported in part by Department of Energy contract DE-AC02-76SF00515.

and z=20.5 mag (Abazajian et al).The ds9 program was used to view these FITS files to determine where there was overlap. The fields from each run almost always line up with fields from an adjacent run, offset by half a field. The overlapping fields and runs are listed in Table 1. In some cases, there are gaps in each camcol in the stripes. For example, in camcol 6 of stripe 34, there is a large gap between fields 302 and 306 of run 2243, overlapping directly with a gap between field 117 of run 2987 and field 11 of run 2989. Additionally, in other cases a field in one run is covered by two or more fields in its overlapping run. This happened most often when one run ended and another began, such as with run 2076 field 172 and run 1895 field 160 in camcol 1 stripe 35. This characteristic was included in the code we developed to analyze the photometric output.

TABLE 1

Run and Field overlap	
Run[1] (35)[2]	Field (35)
1889, 2076	40-122, 40-122
2075, 2076	30-69, 122-162
1889, 2076	162-170, 162-170
2074, 2076	76-78, 170-172
2074, 1895	78-183, 160-265
1896, 1895	46-100, 265-319
1896, 2298	100-131, 61-92
1896, 2304	131-186, 162-216
2326, 2304	84-207, 216-339
2326, 2299	207-313, 198-304
2328, 2299	11-25, 304-318
2328, 2305	25-179, 26-180
Run (34)	Field (34)
2131, 2137	17-128, 47-158
2243, 2137	52-205, 158-310
2243, 2987	206-302, 20-117
2243, 2989	306-411, 11-115
3183, 3225	85-127, 24-67
3184, 3225	11-92, 67-148
3226, 3225	11-226, 148-363

[1]In Run (x, y) and Field (x, y), field x corresponds to run x and field y corresponds to run y

[2]This denotes the stripe number

The photometric analyses were performed using the Sextractor tool (Bertin). Each field contains 2048 pixels in the X direction and 1361 unique pixels in the Y directions. For the two stripes we have examined, approximately 140 pixels of both edges in the X direction (the left and right edges) are covered again in the neighboring strip. The (X,Y) coordinates for each source were then converted into (RA,Dec). We also used the target summary FITS files for each field to perform photometric calibration. (Abazajian et al.). The equations

$$m = -\frac{2.5}{ln10}\left[asinh\left(\frac{f/f_0}{2b}\right) + ln(b)\right]$$

$$\frac{f}{f_0} \longrightarrow \frac{N}{t_{exp}} 10^{0.4(aa+kk\times airmass)}$$

TABLE 2

Asinh Magnitude Softening Parameters	
Band	b
u	1.4×10^{-10}
g	0.9×10^{-10}
r	1.2×10^{-10}
i	1.8×10^{-10}
z	7.4×10^{-10}

from the Second Data Release of the SDSS (Abazajian et al) provided us with an appropriate method to do this conversion. They give the conversion between the SDSS asinh magnitudes, Janskys, and imaging camera counts (N). The quantities aa (photometric zeropoint), kk (extinction term), and $airmass$ were taken from the target summary FITS files, while b can be found for each filter in table 2. T_{exp} is the time exposure for each SDSS pixel, equal to 53.907456 s.

A typical field contains approximately 700 sources in r, of which ∼30 of them lie on one side of the overlap region with its adjacent runs. Photometric measurements are, therefore, available during two epochs. Overlap only occurred between runs on stripes. We found that an overlap on two parallel runs in a single camcol yielded a total area of approximately 1.0' × 8.9'. The observation times of the first fields of each run are listed in Table 1. We detected 48654 sources in filter r, 41118 in filter i, 27939 in filter g, 7028 in filter u, and 24725 in filter z.

We then searched for sources that are detected

in both adjacent fields. This was done simply by matching their coordinates. When the source locations were found to be within 1 arcsecond, they were considered to be the same source. Magnitudes and their statistical uncertainties in each filter and each field were recorded. The final catalogue contains the RA and Dec for each source, 10 magnitude measurements (5 filters and 2 observations) and their uncertainties. Numerous sources were detected in one field but not in the other, primarily due to their faintness. This catalogue contains 24884 sources from Stripe 34 Camcol 6 and 23770 sources from Stripe 35 Camcol 1 and can be used to perform variability studies of the optical sky.

3. Results

We first compute the magnitude differences between the two observations of the sources in each of the 5 filters (Δm_x, where $x = u, g, r, i, z$). The distribution of Δm_x are shown in Figures 1 – 6 for the different magnitude intervals. The shapes of these distributions are modeled as gaussians with 3 free parameters: (1) centroid, (2) gaussian width (sigma), and (3) normalization. We are primarily interested in the gaussian width, which characterizes the amount of variability present in the optical sky. Of interest, is the fact that the measured sigma is larger than expected, by many orders of magnitude. The intrinsic photometric uncertainty for this study was expected to be $\sim$ 0.02 magnitude; as the figures show, the actual width was much larger, and varied according source magnitude (between $v = 19$ to 20 or $v = 21$ to 22 for example) and the different filters. If there had been no magnitude difference detected from one side of the camera column to the other, then the resultant sigma would be equivalent to 0.02 magnitude. Table 4 presents the measured gaussian width for the five filters in the two columns studied. We also noted the correlation of increase in sigma to increase in magnitude; as the sources studied became fainter, their variability grew. Additionally, we found a correlation between the gaussian width and the type of filter. Starting in ultraviolet (u) and going to far infrared (z), one can plainly see that the sigma increases. This means that, for the most part, sources varied more in the far infrared than in the ultraviolet, steadily increasing between the two.

Figures 1 – 6 present evidence for highly variable unusual sources. They are located in the tails of the fitted gaussian. These unusual sources have a change in magnitude ranging from about 0.2 to even $\sim$ 2.5 magnitudes. Their very existence supports the idea that SDSS is a useful tool for finding unusual variable sources.

In searching for variable sources, most found thus far are stars, galaxies, and quasars. This is demonstrated through the color - color plots (g-r vs u-g for two different camcols) in Figures 8 and 9. Ivezic et al. claim that stars will be clumped in the center of the band, while unusual sources will lie scattered around it. In particular, in their paper Variability Studies with SDSS, they state that quasars are expected to be located in the third quadrant of the figure just past the thick band of stars (Ivezic et al. 2004). Moreover, they suggest that sources brighter than quasars lie even farther to the bottom left in this quadrant (Ivezic et al. 2004). The hope is to find unusual sources such as supernovae, black holes, and gamma ray bursts. The fact that no unusual sources have been found yet is not entirely unexpected, as the data in this paper only reflects $\sim$ 1% of the total data available on the SDSS website. At this point, this is only a preliminary presentation of what can be possible with this data, once all of it is studied in this fashion.

Following Ivezic et al., we plot the color - color properties of these sources in the g-r – u-g plane to search for non-stellar sources with unusual color. Figures 7 and 8 show these color difference plots. The thick part of the band is presumably comprised mainly of stars, while outliers to the bottom right dominated by colors of red are mostly unusual galaxies. The method used to find sources in these graphs was based on the fact that stars lie in the central thick band, and other unusual sources lie around it (Ivezic et al. 2004). We looked at $\sim$ 40% of the sources in the fourth quadrant. Upon looking at the actual image of these sources, some were found to be false. By false, we mean that in one epoch, the source is half-off the field. The filter for "flagged" objects did not work 100%, because it was supposed to not include such sources in the final catalogue. This makes an obvious change in detected magnitude, as only $\sim$ half of the photons from the

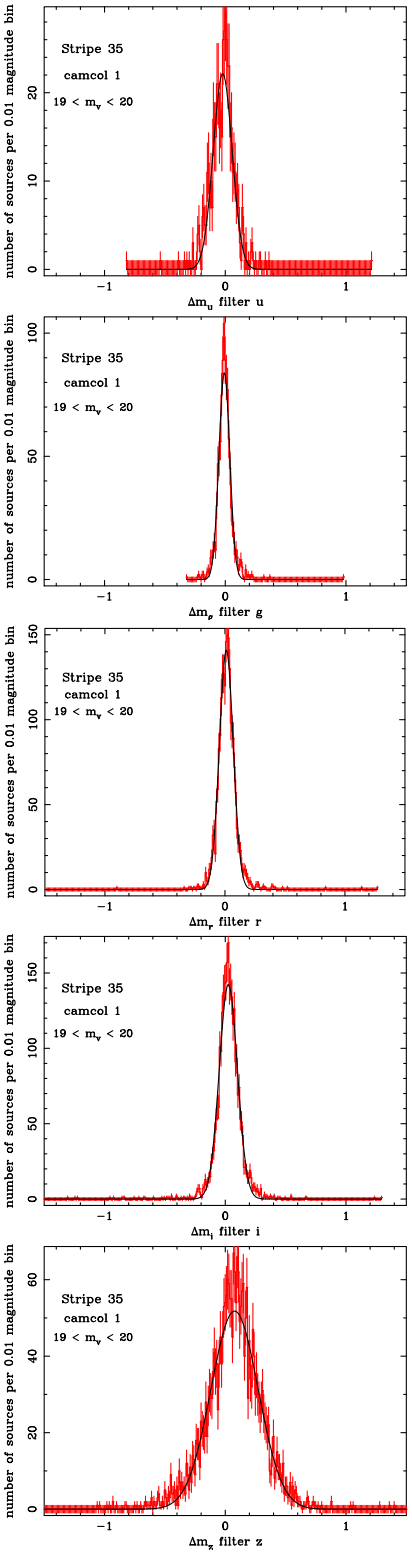

Figure 1. Stripe 35, camcol 1, $19 < v < 20$, done for each filter, ugriz

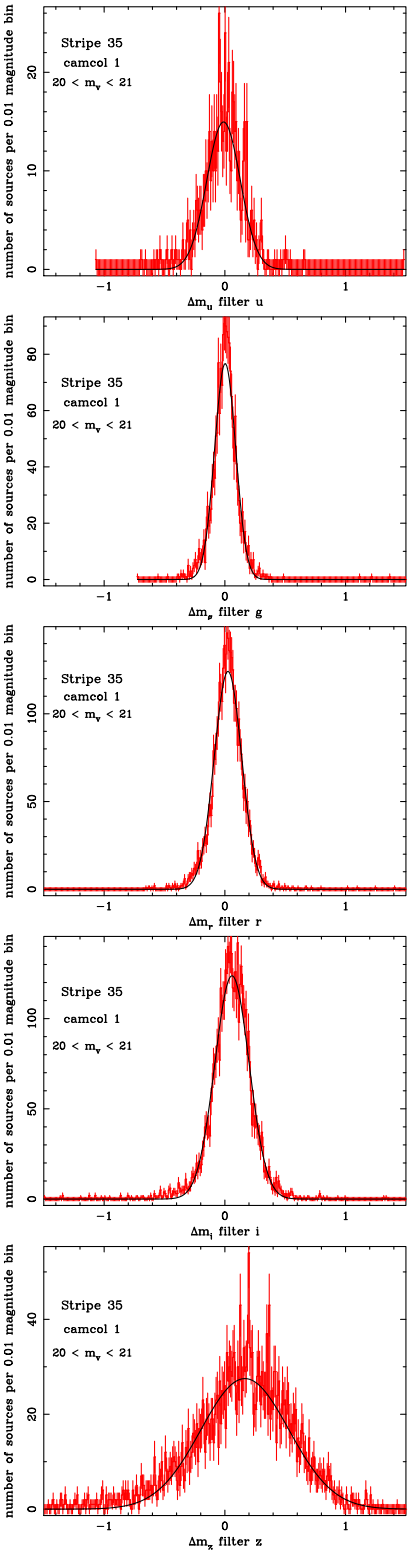

Figure 2. Stripe 35, camcol 1, $20 < v < 21$, done for each filter, ugriz

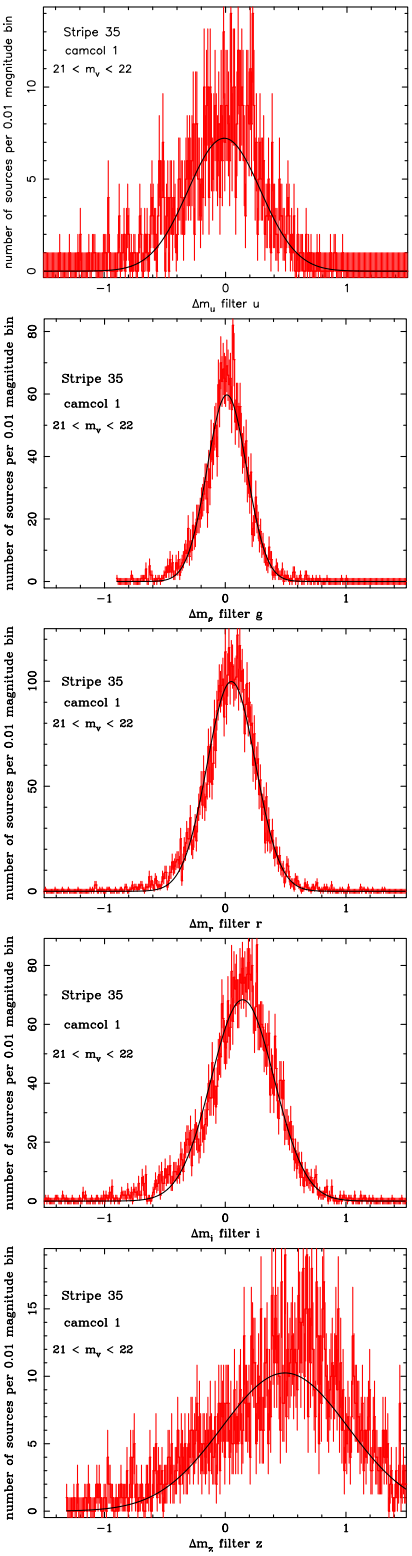

Figure 3. Stripe 35, camcol 1, $21 < v < 22$, done for each filter, ugriz

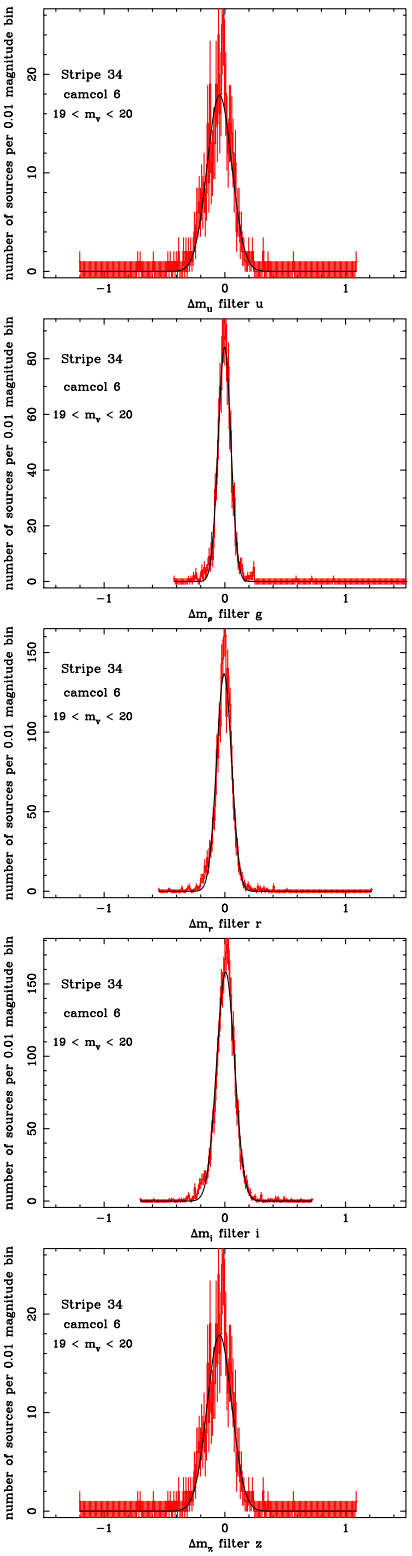

Figure 4. Stripe 34, camcol 6, $19 < v < 20$, done for each filter, ugriz

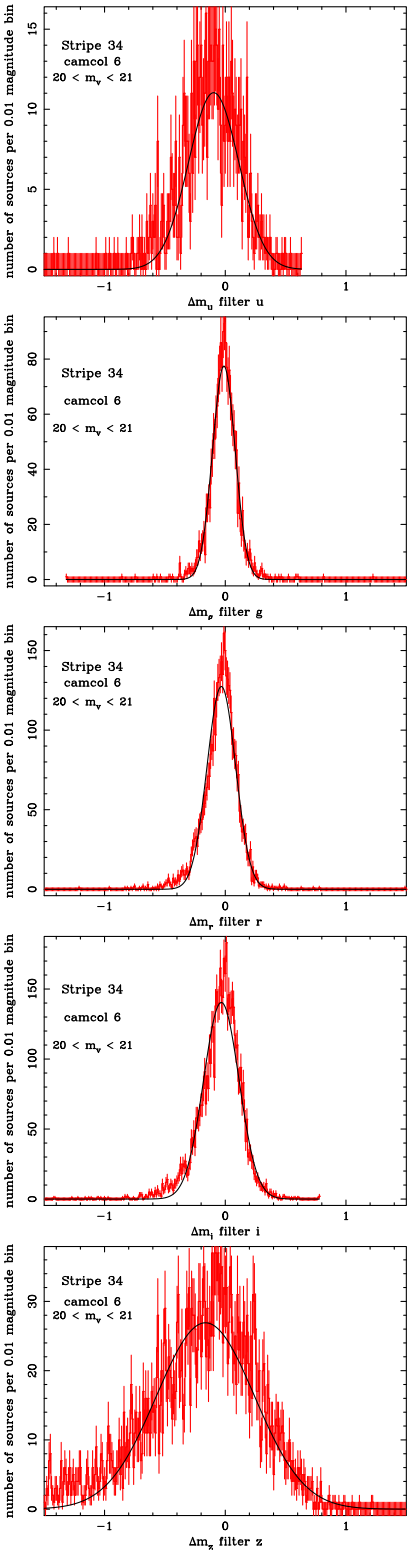

Figure 5. Stripe 34, camcol 6, $20 < v < 21$, done for each filter, ugriz

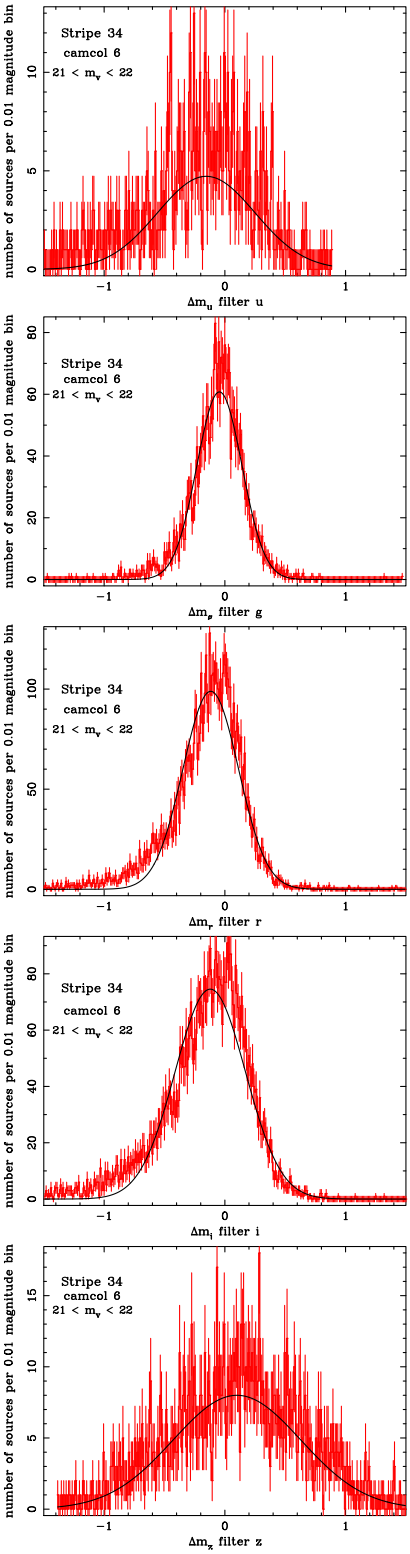

Figure 6. Stripe 34, camcol 6, $21 < v < 22$, done for each filter, ugriz

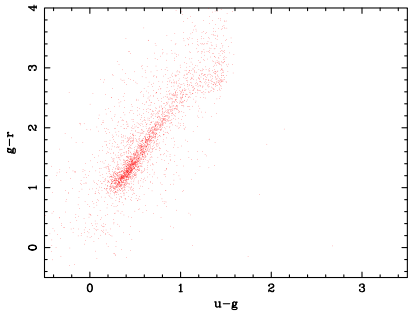

Figure 7. color - color diagram for stripe 35 cam 1

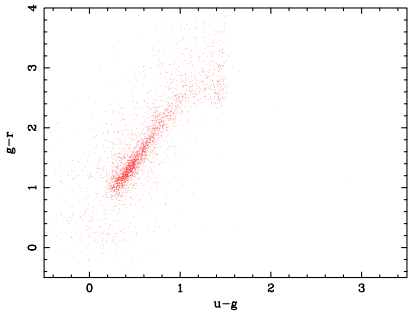

Figure 8. color - color diagram for stripe 34 cam 6

original source were collected in the second epoch. However, this effect is small as it only affects a small fraction of the sources. Moreover, it does not change the gaussian width results, as the majority of the gaussian width is fit for relatively small changes in magnitude. Any false sources would be located on the tails of the fits, not contributing to the sigma.

The catalogue was searched for sources that exhibited large variability $\Delta m_r > 1.5$. Thus far the sources meeting this criteria are either unusual galaxies or false sources (as explained above). At 08:58:54.4824, 49:23:28.68 in Stripe 35, camcol 1 exists an unusual galaxy. This galaxy was located in run 1895, field 186 and run 2074, field 105. Its change in magnitude for all five filters was 0.1291 (r) 0.1002 (i) 0.005 (g) 1.576 (u) 0.0158 (z). As of yet in our study, no supernovae or other unusual sources have been found. However, the calculated statistical probability that we will find one in only two camcols is low. Only one unusual source per three camera columns is expected. There is a high probability that we will find an unusual source in one camera column out of three studied. To date, we have only looked at two camera columns, so it is highly probable that an unusual source will be found in a third camera column.

4. Conclusion

The Sloan Digital Sky Survey is a crucial aid to researching optically faint variable sources. It serves as a mini-survey of future projects such as LSST and SNAP. Resulting from research done thus far, we know that the sky's variability uncertainty is far greater than was previously thought. We have measured the magnitude difference distribution of the optical sky. These distributions demonstrate that there is a significant level of optical variability, and that this variability is a funtion of source magnitude and color. The intrinsic photometric uncertainty is 0.02 magnitude, yet from only two camera columns it was found to be greater than 0.04 for some, and as large as 0.5 for others. This characterizes the amount of variability present in the optical sky. Additionally, we found color - color diagrams to be a useful approach to pinpointing unusual sources. Based on the knowledge that stars lie in the central band (Ivezic et al. 2004), we looked at a majority of the individual surrounding sources in the ds9 FITS image viewer. While some sources were found to be false, a few unusual variable galaxies were found. The problem with false sources can solved with further study by simply limiting the region of study within the overlapping fields. Most importantly, this shows that using SDSS as a preliminary search for variable sources is a productive application, as nearly 50,000 sources were found in only two camera columns. This comprises less than 1% of the total publicly available data from SDSS. No supernovae were found thus far in two camera columns. However, based on the statistical probability that one will be found in three camera columns, this result is not surprising. One can expect to find a supernova source in a third camera column. Once all the data is studied, there is a high probability of finding unusual sources that warrant followup studies.

5. References

Abazajian, K., et al. 2004, ArXiv Astrophysics e-prints, astro-ph/0403325v1
E. Bertin. SExtractor, 2.1.3 edition, 1999. Newsgroup URL:sextractor@iap.fr.
Ivezic, Z., et al. 2003, ArXiv Astrophysics e-prints, astro-ph/0301400
Ivezic, A., et al. 2004, ArXiv Astrophysics e-prints, astro-ph/0404487
Szabo, Gy. M., Ivezic, Z., Juric, M., Lupton, R., & Kiss, L. L. 2004, MNRAS, 348, 987
www.sdss.org.dr2

TABLE 3

Times and Dates for each Run[1]		
Run (35)	Date	time (TAI, HMS)
1889[2]	11/26/2000	08:32:21.55
2076	01/26/2001	08:05:10.70
2075	01/26/2001	06:59:53.43
1889[3]	11/26/2000	09:45:12.07
1895	11/27/2000	07:54:28.87
1896	11/27/2000	10:29:55.13
2298	05/17/2001	04:04:14.69
2074	01/26/2001	05:13:12.26
2326	05/23/2001	04:38:10.00
2304	05/18/2001	04:15:48.21

[1]taken from first field
[2]starting at field 40
[3]starting at field 162

TABLE 4

gaussian width σ					
Magnitude[1]	u	g	r	i	z
$19 \leq v \leq 20$	0.079 ± 0.006	0.045 ± 0.003	0.061 ± 0.002	0.075 ± 0.002	0.196 ± 0.005
$20 \leq v \leq 21$	0.014 ± 0.009	0.087 ± 0.004	0.114 ± 0.003	0.142 ± 0.003	0.3655 ± 0.0105
$21 \leq v \leq 22$	0.299 ± 0.025	0.164 ± 0.005	0.201 ± 0.004	0.261 ± 0.006	0.523 ± 0.023
Magnitude[2]	u	g	r	i	z
$19 \leq v \leq 20$	0.102 ± 0.007	0.052 ± 0.002	0.068 ± 0.002	0.074 ± 0.002	0.102 ± 0.007
$20 \leq v \leq 21$	0.208 ± 0.140	0.094 ± 0.003	0.118 ± 0.003	0.147 ± 0.003	0.406 ± 0.012
$21 \leq v \leq 22$	0.405 ± 0.046	0.187 ± 0.005	0.231 ± 0.005	0.291 ± 0.006	0.530 ± 0.029

[1]for Stripe 35 camcol 1
[2]for Stripe 34 camcol 6