

## THE SECOND DATA RELEASE OF THE SLOAN DIGITAL SKY SURVEY

KEVORK ABAZJIAN,<sup>1</sup> JENNIFER K. ADELMAN-McCARTHY,<sup>2</sup> MARCEL A. AGÜEROS,<sup>3</sup> SAHAR S. ALLAM,<sup>4</sup> KURT, S. J. ANDERSON,<sup>4,5</sup>  
 SCOTT F. ANDERSON,<sup>3</sup> JAMES ANNIS,<sup>2</sup> NETA A. BAHCALL,<sup>6</sup> IVAN K. BALDRY,<sup>7</sup> STEVEN BASTIAN,<sup>2</sup> ANDREAS BERLIND,<sup>8,9,10</sup>  
 MARIANGELA BERNARDI,<sup>11</sup> MICHAEL R. BLANTON,<sup>10</sup> JOHN J. BOCHANSKI, JR.,<sup>3</sup> WILLIAM N. BOROSKI,<sup>2</sup> JOHN W. BRIGGS,<sup>8</sup> J. BRINKMANN,<sup>5</sup>  
 ROBERT J. BRUNNER,<sup>12</sup> TAMÁS BUDAVÁRI,<sup>7</sup> LARRY N. CAREY,<sup>3</sup> SAMUEL CARLILES,<sup>7</sup> FRANCISCO J. CASTANDER,<sup>13</sup> A. J. CONNOLLY,<sup>14</sup>  
 ISTVÁN CSABAI,<sup>7,15</sup> MAMORU DOI,<sup>16</sup> FENG DONG,<sup>6</sup> DANIEL J. EISENSTEIN,<sup>17</sup> MICHAEL L. EVANS,<sup>3</sup> XIAOHUI FAN,<sup>17</sup> DOUGLAS P. FINKBEINER,<sup>6</sup>  
 SCOTT D. FRIEDMAN,<sup>18</sup> JOSHUA A. FRIEMAN,<sup>2,8,9</sup> MASATAKA FUKUGITA,<sup>19</sup> ROY R. GAL,<sup>7</sup> BRUCE GILLESPIE,<sup>5</sup> KARL GLAZEBROOK,<sup>7</sup>  
 JIM GRAY,<sup>20</sup> EVA K. GREBEL,<sup>21</sup> JAMES E. GUNN,<sup>6</sup> VIJAY K. GURBANI,<sup>22,2</sup> PATRICK B. HALL,<sup>6</sup> MASARU HAMABE,<sup>23</sup> FREDERICK H. HARRIS,<sup>24</sup>  
 HUGH C. HARRIS,<sup>24</sup> MICHAEL HARVANEK,<sup>5</sup> TIMOTHY M. HECKMAN,<sup>7</sup> JOHN S. HENDRY,<sup>2</sup> GREGORY S. HENNESSY,<sup>25</sup> ROBERT B. HINDSLEY,<sup>26</sup>  
 CRAIG J. HOGAN,<sup>3</sup> DAVID W. HOGG,<sup>10</sup> DONALD J. HOLMGREN,<sup>2</sup> SHIN-ICHI ICHIKAWA,<sup>27</sup> TAKASHI ICHIKAWA,<sup>28</sup> ŽELJKO IVEZIĆ,<sup>6</sup>  
 SEBASTIAN JESTER,<sup>2</sup> DAVID E. JOHNSTON,<sup>6,8,9</sup> ANDERS M. JORGENSEN,<sup>29</sup> STEPHEN M. KENT,<sup>2</sup> S. J. KLEINMAN,<sup>5</sup> G. R. KNAPP,<sup>6</sup>  
 ALEXEI YU. KNAZEV,<sup>21</sup> RICHARD G. KRON,<sup>2,8</sup> JUREK KRZESINSKI,<sup>5,30</sup> PETER Z. KUNSZT,<sup>7,31</sup> NICKOLAI KUROPATKIN,<sup>2</sup> DONALD Q. LAMB,<sup>8,32</sup>  
 HUBERT LAMPEITL,<sup>2</sup> BRIAN C. LEE,<sup>33</sup> R. FRENCH LEGER,<sup>2</sup> NOLAN LI,<sup>7</sup> HUAN LIN,<sup>2</sup> YEONG-SHANG LOH,<sup>6,34</sup> DANIEL C. LONG,<sup>5</sup>  
 JON LOVEDAY,<sup>35</sup> ROBERT H. LUPTON,<sup>6</sup> TANU MALIK,<sup>7</sup> BRUCE MARGON,<sup>18</sup> TAKAHIKO MATSUBARA,<sup>36</sup> PEREGRINE M. McGEHEE,<sup>4,37</sup>  
 TIMOTHY A. MCKAY,<sup>38</sup> AVERY MEIKSIN,<sup>39</sup> JEFFREY A. MUNN,<sup>24</sup> REIKO NAKAJIMA,<sup>40</sup> THOMAS NASH,<sup>2</sup> ERIC H. NEILSEN, JR.,<sup>2</sup>  
 HEIDI JO NEWBERG,<sup>41</sup> PETER R. NEWMAN,<sup>5</sup> ROBERT C. NICHOL,<sup>11</sup> TOM NICHINSKI,<sup>2,42</sup> MARIA NIETO-SANTISTEBAN,<sup>7</sup> ATSUKO NITTA,<sup>5</sup>  
 SADANORI OKAMURA,<sup>43</sup> WILLIAM O’MULLANE,<sup>7</sup> JEREMIAH P. OSTRICKER,<sup>6</sup> RUSSELL OWEN,<sup>3</sup> NIKHIL PADMANABHAN,<sup>44</sup> JOHN PEOPLES,<sup>2</sup>  
 JEFFREY R. PIER,<sup>24</sup> ADRIAN C. POPE,<sup>7</sup> THOMAS R. QUINN,<sup>3</sup> GORDON T. RICHARDS,<sup>6</sup> MICHAEL W. RICHMOND,<sup>45</sup> HANS-WALTER RIX,<sup>21</sup>  
 CONSTANCE M. ROCKOSI,<sup>3</sup> DAVID J. SCHLEGEL,<sup>6</sup> DONALD P. SCHNEIDER,<sup>46</sup> RYAN SCRANTON,<sup>14</sup> MAKI SEKIGUCHI,<sup>47</sup> UROS SELJAK,<sup>6</sup>  
 GARY SERGEY,<sup>2</sup> BRANIMIR SESAR,<sup>48</sup> ERIN SHELDON,<sup>8,9</sup> KAZU SHIMASAKU,<sup>43</sup> WALTER A. SIEGMUND,<sup>49</sup> NICOLE M. SILVESTRI,<sup>3</sup>  
 J. ALLYN SMITH,<sup>29,50</sup> VERNESA SMOLČIĆ,<sup>48</sup> STEPHANIE A. SNEDDEN,<sup>5</sup> ALBERT STEBBINS,<sup>2</sup> CHRIS STOUGHTON,<sup>2</sup> MICHAEL A. STRAUSS,<sup>6</sup>  
 MARK SUBBARAO,<sup>51,8</sup> ALEXANDER S. SZALAY,<sup>7</sup> ISTVÁN SZAPUDI,<sup>49</sup> PAULA SZKODY,<sup>3</sup> GYULA P. SZOKOLY,<sup>52</sup> MAX TEGMARK,<sup>40</sup> LUIS TEODORO,<sup>1</sup>  
 ANIRUDDHA R. THAKAR,<sup>7</sup> CHRISTY TREMONTI,<sup>17</sup> DOUGLAS L. TUCKER,<sup>2</sup> ALAN UOMOTO,<sup>53,7</sup> DANIEL E. VANDEN BERK,<sup>14</sup> JAN VANDENBERG,<sup>7</sup>  
 MICHAEL S. VOGELY,<sup>54</sup> WOLFGANG VOGES,<sup>52</sup> NICOLE P. VOGT,<sup>4</sup> LUCIANNE M. WALKOWICZ,<sup>3</sup> SHU-I WANG,<sup>8</sup> DAVID H. WEINBERG,<sup>55</sup>  
 ANDREW A. WEST,<sup>3</sup> SIMON D. M. WHITE,<sup>56</sup> BRIAN C. WILHITE,<sup>8</sup> YONGZHONG XU,<sup>1</sup> BRIAN YANNY,<sup>2</sup> NAOKI YASUDA,<sup>19</sup> CHING-WA YIP,<sup>14</sup>  
 D. R. YOCUM,<sup>2</sup> DONALD G. YORK,<sup>8,32</sup> IDIT ZEHAVI,<sup>17</sup> STEFANO ZIBETTI,<sup>56</sup> AND DANIEL B. ZUCKER<sup>21</sup>

Received 2004 March 8; accepted 2004 March 16

## ABSTRACT

The Sloan Digital Sky Survey (SDSS) has validated and made publicly available its Second Data Release. This data release consists of 3324 deg<sup>2</sup> of five-band (*ugriz*) imaging data with photometry for over 88 million unique objects, 367,360 spectra of galaxies, quasars, stars, and calibrating blank sky patches selected over 2627 deg<sup>2</sup> of this area, and tables of measured parameters from these data. The imaging data reach a depth of  $r \approx 22.2$  (95% completeness limit for point sources) and are photometrically and astrometrically calibrated to 2% rms and 100 mas rms per coordinate, respectively. The imaging data have all been processed through a new version of the SDSS imaging pipeline, in which the most important improvement since the last data release is fixing an error in the model fits to each object. The result is that model magnitudes are now a good proxy for point-spread function magnitudes for point sources, and Petrosian magnitudes for extended sources. The spectroscopy extends from 3800 to 9200 Å at a resolution of 2000. The spectroscopic software now repairs a systematic error in the radial velocities of certain types of stars and has substantially improved spectrophotometry. All data included in the SDSS Early Data Release and First Data Release are reprocessed with the improved pipelines and included in the Second Data Release. Further characteristics of the data are described, as are the data products themselves and the tools for accessing them.

*Key words:* atlases — catalogs — surveys

<sup>1</sup> Theoretical Division, MS B285, Los Alamos National Laboratory, Los Alamos, NM 87545.

<sup>2</sup> Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, IL 60510.

<sup>3</sup> Department of Astronomy, University of Washington, Box 351580, Seattle, WA 98195.

<sup>4</sup> Department of Astronomy, MSC 4500, New Mexico State University, P.O. Box 30001, Las Cruces, NM 88003.

<sup>5</sup> Apache Point Observatory, P.O. Box 59, Sunspot, NM 88349.

<sup>6</sup> Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544.

<sup>7</sup> Center for Astrophysical Sciences, Department of Physics and Astronomy, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218.

<sup>8</sup> Department of Astronomy and Astrophysics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637.

<sup>9</sup> Center for Cosmological Physics, The University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637.

<sup>10</sup> Center for Cosmology and Particle Physics, Department of Physics, New York University, 4 Washington Place, New York, NY 10003.

<sup>11</sup> Department of Physics, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213.

<sup>12</sup> Department of Astronomy, University of Illinois, 1002 West Green Street, Urbana, IL 61801.

<sup>13</sup> Institut d’Estudis Espacials de Catalunya/CSIC, Gran Capitá 2-4, E-08034 Barcelona, Spain.

<sup>14</sup> Department of Physics and Astronomy, University of Pittsburgh, 3941 O’Hara Street, Pittsburgh, PA 15260.

<sup>15</sup> Department of Physics of Complex Systems, Eötvös Loránd University, Pf. 32, H-1518 Budapest, Hungary.

<sup>16</sup> Institute of Astronomy and Research Center for the Early Universe, School of Science, University of Tokyo, 2-21-1 Osawa, Mitaka, Tokyo 181-0015, Japan.

<sup>17</sup> Steward Observatory, 933 North Cherry Avenue, Tucson, AZ 85721.

## Report Documentation Page

*Form Approved*  
*OMB No. 0704-0188*

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE <b>16 MAR 2004</b>	2. REPORT TYPE <b>N/A</b>	3. DATES COVERED <b>-</b>	
4. TITLE AND SUBTITLE <b>The Second Data Release Of The Sloan Digital Sky Survey</b>		5a. CONTRACT NUMBER	
		5b. GRANT NUMBER	
		5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)		5d. PROJECT NUMBER	
		5e. TASK NUMBER	
		5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>US Naval Observatory, Flagstaff Station, P.O. Box 1149, Flagstaff, AZ 86002-1149</b>		8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)	
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release, distribution unlimited</b>			
13. SUPPLEMENTARY NOTES			
14. ABSTRACT <b>The Sloan Digital Sky Survey (SDSS) has validated and made publicly available its Second Data Release. This data release consists of 3324 deg<sup>2</sup> of five-band (ugriz) imaging data with photometry for over 88 million unique objects, 367,360 spectra of galaxies, quasars, stars, and calibrating blank sky patches selected over 2627 deg<sup>2</sup> of this area, and tables of measured parameters from these data. The imaging data reach a depth of r 22:2 (95% completeness limit for point sources) and are photometrically and astrometrically calibrated to 2% rms and 100 mas rms per coordinate, respectively. The imaging data have all been processed through a new version of the SDSS imaging pipeline, in which the most important improvement since the last data release is fixing an error in the model fits to each object. The result is that model magnitudes are now a good proxy for point-spread function magnitudes for point sources, and Petrosian magnitudes for extended sources. The spectroscopy extends from 3800 to 9200 Å at a resolution of 2000. The spectroscopic software now repairs a systematic error in the radial velocities of certain types of stars and has substantially improved spectrophotometry. All data included in the SDSS Early Data Release and First Data Release are reprocessed with the improved pipelines and included in the Second Data Release. Further characteristics of the data are described, as are the data products themselves and the tools for accessing them.</b>			
15. SUBJECT TERMS <b>atlases; catalogs; surveys</b>			
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT <b>UU</b>
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>	
			18. NUMBER OF PAGES <b>11</b>
			19a. NAME OF RESPONSIBLE PERSON

## 1. THE SLOAN DIGITAL SKY SURVEY

The Sloan Digital Sky Survey (SDSS; York et al. 2000) is an imaging and spectroscopic survey of the high Galactic latitude sky visible from the northern hemisphere. The principal survey goals are to measure the large-scale distribution of galaxies and quasars and to produce an imaging and spectroscopic legacy for the astronomical community. The SDSS data have been used in well over 200 refereed papers to date on subjects ranging from the colors of asteroids (Ivezić et al. 2002), to magnetic white dwarfs (Schmidt et al. 2003), to structures in the Galactic halo (Newberg et al. 2003), to the star formation history of galaxies (Kauffmann et al. 2003), to Type II quasars (Zakamska et al. 2003), to the large-scale distribution of galaxies (Pope et al. 2004; Tegmark et al. 2004). The survey uses a dedicated 2.5 m telescope with a  $3^\circ$  field of view at Apache Point Observatory, New Mexico. A 120 megapixel camera (Gunn et al. 1998) images in five broad bands ( $u$ ,  $g$ ,  $r$ ,  $i$ , and  $z$ ; Fukugita et al. 1996; Stoughton et al. 2002) on clear moonless nights of good seeing. These data are photometrically calibrated using an auxiliary 20 inch (0.5 m) telescope with a  $40' \times 40'$  imager, which determines the photometricity of each night (Hogg et al. 2001) and measures the extinction and photometric zero point using a network of standard stars (Smith et al. 2002).

The imaging data are processed through a series of pipelines that locate and measure the properties of all detected objects (Lupton et al. 2001) and carry out photometric and astrometric calibration (Pier et al. 2003). From the resulting catalogs of objects, complete catalogs of galaxies (Eisenstein

et al. 2001; Strauss et al. 2002) and quasar candidates (Richards et al. 2002) are selected for spectroscopic follow-up, and are assigned to spectroscopic tiles of diameter  $3^\circ$  (Blanton et al. 2003). Spectroscopy is performed on nights with moonlight, mild cloud cover, and/or poor seeing using a pair of double spectrographs with coverage from 3800 to 9200 Å, and a resolution  $\lambda/\Delta\lambda$  of roughly 2000. A plug plate for each tile holds 640 optical fibers of  $3''$  entrance aperture which feed the spectrographs, together with 11 coherent fiber bundles to image guide stars. Because of the diameter of the cladding holding the optical fibers, spectroscopy cannot be carried out for objects separated by less than  $55''$  on a given plate.

## 2. THE SECOND DATA RELEASE

A high-level overview of the SDSS may be found in York et al. (2000), while many of the details of the software and data products may be found in Stoughton et al. (2002). The latter paper also describes our Early Data Release (EDR), which consisted mostly of data taken during our commissioning period. The First Data Release (DR1) was made public in 2003 April; Abazajian et al. (2003) describe these data and give further details and updates on the quality of the data and its processing. The current paper describes the Second Data Release of the SDSS (DR2), which was made available to the public on 2004 March 15. The properties of DR2 are summarized in Table 1. The DR2 footprint is defined by all non-repeating survey-quality imaging runs within the a priori defined elliptical survey area in the northern Galactic cap and three stripes in the southern Galactic cap (York et al. 2000) obtained prior to 2002 July 1, as well as the spectroscopy

- 
- <sup>18</sup> Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218.  
<sup>19</sup> Institute for Cosmic Ray Research, University of Tokyo, 5-1-5 Kashiwa, Kashiwa City, Chiba 277-8582, Japan.  
<sup>20</sup> Microsoft Research, 455 Market Street, Suite 1690, San Francisco, CA 94105.  
<sup>21</sup> Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany.  
<sup>22</sup> Lucent Technologies, 2000 Lucent Lane, Naperville, IL 60566.  
<sup>23</sup> Department of Mathematical and Physical Sciences, Japan Women's University, 2-8-1 Mejirodai, Bunkyo, Tokyo 112-8681, Japan.  
<sup>24</sup> US Naval Observatory, Flagstaff Station, P.O. Box 1149, Flagstaff, AZ 86002-1149.  
<sup>25</sup> US Naval Observatory, 3540 Massachusetts Avenue NW, Washington, DC 20392.  
<sup>26</sup> Code 7215, Remote Sensing Division, Naval Research Laboratory, 4555 Overlook Avenue SW, Washington, DC 20392.  
<sup>27</sup> National Astronomical Observatory, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan.  
<sup>28</sup> Astronomical Institute, Tohoku University, Aramaki, Aoba, Sendai 980-8578, Japan.  
<sup>29</sup> ISR-4, MS D448, Los Alamos National Laboratory, P.O. Box 1663, Los Alamos, NM 87545.  
<sup>30</sup> Obserwatorium Astronomiczne na Suhorze, Akademia Pedagogiczna w Krakowie, ulica Podchorążych 2, PL-30-084 Kraców, Poland.  
<sup>31</sup> CERN, IT Division, 1211 Geneva 23, Switzerland.  
<sup>32</sup> Enrico Fermi Institute, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637.  
<sup>33</sup> Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley CA 94720-8160.  
<sup>34</sup> Center for Astrophysics and Space Astronomy, University of Colorado, Boulder, CO 80309.  
<sup>35</sup> Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QJ, UK.  
<sup>36</sup> Department of Physics and Astrophysics, Nagoya University, Chikusa, Nagoya 464-8602, Japan.  
<sup>37</sup> SNS-4, MS H820, Los Alamos National Laboratory, P.O. Box 1663, Los Alamos, NM 87545.  
<sup>38</sup> Department of Physics, University of Michigan, 500 East University Avenue, Ann Arbor, MI 48109.  
<sup>39</sup> Institute for Astronomy, Royal Observatory, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK.  
<sup>40</sup> Department of Physics, University of Pennsylvania, Philadelphia, PA 19104.  
<sup>41</sup> Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, 110 Eighth Street, Troy, NY 12180.  
<sup>42</sup> CMC Electronics Aurora, 43W752 Route 30, Sugar Grove, IL 60554.  
<sup>43</sup> Department of Astronomy and Research Center for the Early Universe, University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan.  
<sup>44</sup> Joseph Henry Laboratories, Princeton University, Princeton, NJ 08544.  
<sup>45</sup> Department of Physics, Rochester Institute of Technology, 84 Lomb Memorial Drive, Rochester, NY 14623-5603.  
<sup>46</sup> Department of Astronomy and Astrophysics, 525 Davey Laboratory, Pennsylvania State University, University Park, PA 16802.  
<sup>47</sup> Japan Participation Group, c/o Institute for Cosmic Ray Research, University of Tokyo, 5-1-5 Kashiwa, Kashiwa City, Chiba 277-8582, Japan.  
<sup>48</sup> University of Zagreb, Department of Physics, Bijenička cesta 32, 10000 Zagreb, Croatia.  
<sup>49</sup> Institute for Astronomy, 2680 Woodlawn Road, Honolulu, HI 96822.  
<sup>50</sup> Department of Physics and Astronomy, University of Wyoming, Laramie, WY 82071.  
<sup>51</sup> Adler Planetarium and Astronomy Museum, 1300 Lake Shore Drive, Chicago, IL 60605.  
<sup>52</sup> Max-Planck-Institut für extraterrestrische Physik, Giessenbachstrasse 1, D-85741 Garching, Germany.  
<sup>53</sup> Observatories of the Carnegie Institution of Washington, 813 Santa Barbara Street, Pasadena, CA 91101.  
<sup>54</sup> Department of Physics, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104.  
<sup>55</sup> Department of Astronomy, Ohio State University, 140 West 8th Avenue, Columbus, OH 43210.  
<sup>56</sup> Max Planck Institut für Astrophysik, Postfach 1, D-85748 Garching, Germany.

TABLE 1  
CHARACTERISTICS OF THE SDSS SECOND DATA RELEASE

Quantity	Value
Imaging	
Footprint area.....	3324 deg <sup>2</sup>
Imaging catalog.....	88 million unique objects
Magnitude limits: <sup>a</sup>	
<i>u</i> .....	22.0
<i>g</i> .....	22.2
<i>r</i> .....	22.2
<i>i</i> .....	21.3
<i>z</i> .....	20.5
Median PSF width rms photometric calibration errors:	1"4 in <i>r</i>
<i>r</i> .....	2%
<i>u-g</i> .....	3%
<i>g-r</i> .....	2%
<i>r-i</i> .....	2%
<i>i-z</i> .....	3%
Astrometry.....	<0"1 rms absolute per coordinate
Spectroscopy	
Footprint area.....	2627 deg <sup>2</sup>
Wavelength coverage.....	3800–9200 Å
Resolution.....	1800–2100
S/N.....	>4 per pixel at <i>g</i> = 20.2
Wavelength calibration.....	<5 km s <sup>-1</sup>
Redshift accuracy.....	30 km s <sup>-1</sup> rms for main galaxies
Number of spectra.....	367,360

<sup>a</sup> 95% completeness for point sources in typical seeing; 50% completeness numbers are typically 0.4 mag fainter (Abazajian et al. 2003).

associated with that area obtained before that date. In fact, 34 deg<sup>2</sup> of DR2 imaging data in the northern Galactic cap lie outside this ellipse. While the DR2 scans do not repeat a given area of sky, they do overlap to some extent, and the data in the overlaps are included in DR2 as well. The DR2 includes reprocessing of all data included in DR1, and those data in EDR that pass our data-quality criteria for the official survey.

The sky coverage of the imaging and spectroscopic data that make up DR2 is given in Figure 1. The effective areas of the two are 3324 and 2627 deg<sup>2</sup>, respectively. The natural unit of imaging data is a *run*; the DR2 contains data from 105 runs in the *best* database, and 105 runs in the *target* database (*best* and *target* are defined in § 3). Similarly, the natural unit of spectroscopic data is a *plate* of 640 spectra each (of which 32 are devoted to background sky determination); the DR2 contains data from 574 plates.

The DR2 data are available via links off the Web site, which also includes extensive technical information about the SDSS data, and should be checked for errata and caveats to the data.<sup>57</sup> The types of data that are available are described in detail by Stoughton et al. (2002) and Abazajian et al. (2003), and on the above Web site. There are two principal ways to access the data. The first, the Data Archive Server (DAS), allows one to download the FITS files containing the imaging and spectroscopic catalogs, the images themselves, and the spectra. This is the appropriate place to go to download large quantities of data in bulk. The second option, the Catalog Archive Server (CAS), allows one to perform database queries

by object attributes and to obtain finding charts of given regions of sky and plots of the spectra. The CAS also provides pointers to the survey images and spectra in FITS format.

Stoughton et al. (2002) describe the SDSS data in detail. We do not repeat that description here, but put emphasis on changes since DR1 and new-found problems in imaging (§ 3), spectroscopy (§ 4), and target selection (§ 5). An Appendix describes the conversion between magnitudes, fluxes, and counts in the imaging data.

### 3. THE SDSS IMAGING DATA

The SDSS imaging pipelines have evolved as the survey has progressed, leading to continual improvements in the measured quantities. However, this means that much of the spectroscopic targeting is based on reductions carried out with old versions of the imaging pipelines. For this reason, we release two versions of the imaging data for each area of the sky. The *target* version is that which was used for defining spectroscopic targets of a given region, while the *best* version uses the latest version of the imaging pipelines. In some cases, improved data (e.g., with better seeing) were taken in a given region of sky after targeting was done, in which case the *target* and *best* data are independent runs. A total of 183 deg<sup>2</sup> of sky are different runs between *target* and *best*, the majority along the equatorial stripe in the autumn sky.

The quality of the imaging data is described by Stoughton et al. (2002) and Abazajian et al. (2003); the distribution of seeing, sky brightness, and depth for the DR2 data are no different than for DR1, nor are the accuracies of the astrometric and photometric calibrations. The *best* reductions do, however, have some substantial improvements over what

<sup>57</sup> See <http://www.sdss.org/dr2>.

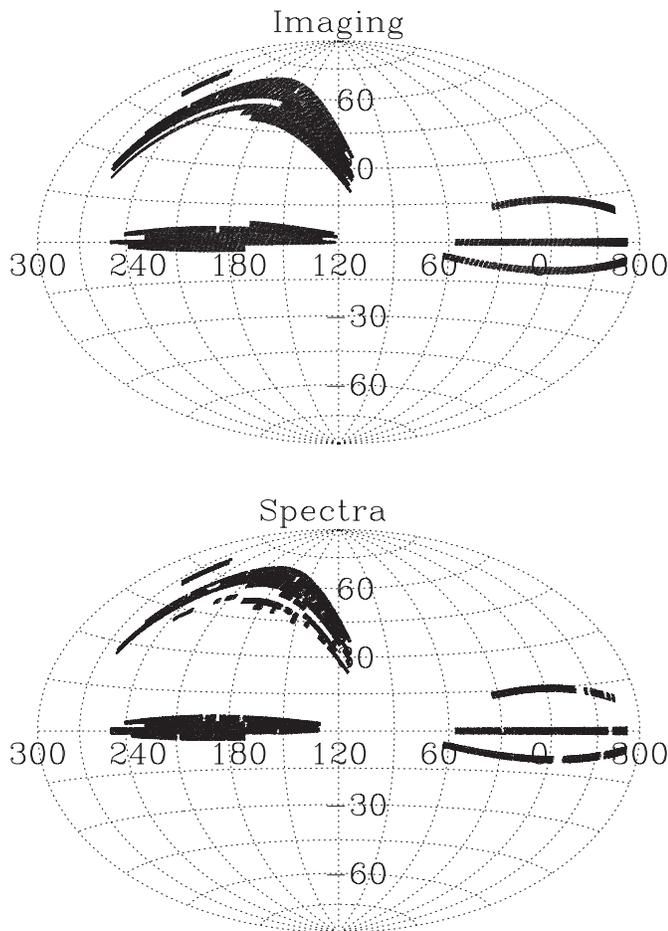


FIG. 1.—Distribution on the sky of the imaging scans and spectroscopic plates included in DR2. This is an Aitoff projection in equatorial coordinates. The total sky area covered by the imaging is 3324 deg<sup>2</sup>, and by the spectroscopy is 2627 deg<sup>2</sup>.

was included in DR1. We now describe these improvements, and give further caveats of problems that have since come to light.

### 3.1. Model Magnitudes

The computation of model magnitudes in the DR1 and EDR processing had a serious bug. The image of each object detected in any of the five bands in the imaging data is fitted to de Vaucouleurs ( $I(\theta) \propto \exp[-(\theta/\theta_0)^{1/4}]$ ) and exponential ( $I(\theta) \propto \exp[-\theta/\theta_0]$ ) radial profiles of arbitrary axis ratio and inclination, convolved with the local point-spread function (PSF). However, these fits used an incorrect model of the PSF, which caused systematic errors in the fit parameters, especially for objects of small scale size  $\theta_0$  (i.e., close to a PSF). For each of these model fits, the code determined an aperture correction to force the exponential and de Vaucouleurs magnitudes to equal the PSF magnitude for stars; this correction was then applied to *all* objects. Because of this software error, this aperture correction was large, 0.2 mag. Thus, the model magnitudes of galaxies were systematically too large by typically 0.2 mag. In the mean, exponential scale lengths  $r_{\text{exp}}$  were overestimated by  $\sim 0.1-0.2$  for objects with  $r_{\text{exp}} < 6''$ , while de Vaucouleurs scale lengths  $r_{\text{deV}}$  were overestimated by a factor 1.25 for large objects, and as much as a factor of 2 for objects with  $r_{\text{deV}} < 0.15$ .

This error has been fixed in the latest version of the pipeline and has been extensively tested with simulations. The code now also takes the best-fit exponential and de Vaucouleurs fits in each band and asks for the linear combination of the two that best fits the image. The coefficient (clipped between zero and one) of the de Vaucouleurs term is stored in the quantity `fracDeV`.<sup>58</sup> This allows us to define a *composite flux*:

$$F_{\text{composite}} = \text{fracDeV} \times F_{\text{deV}} + (1 - \text{fracDeV}) \times F_{\text{exp}}, \quad (1)$$

where  $F_{\text{deV}}$  and  $F_{\text{exp}}$  are the de Vaucouleurs and exponential fluxes (*not* magnitudes) of the object in question. The magnitude derived from  $F_{\text{composite}}$  is referred to below as the `cmodel` magnitude (as distinct from the `model` magnitude, based on the better-fitting of the exponential and de Vaucouleurs models in the *r* band; see Stoughton et al. 2002).

With these changes in place, there is now excellent agreement between `cmodel` and Petrosian magnitudes of galaxies, and `cmodel` and PSF magnitudes of stars (Fig. 2). The `cmodel` and Petrosian magnitudes are not expected to be identical, of course; as Strauss et al. (2002) describe, the Petrosian aperture excludes the outer parts of galaxy profiles, especially for elliptical galaxies. As a consequence, there is an offset of 0.05–0.1 mag between `cmodel` and Petrosian magnitudes of bright galaxies, depending on the photometric bandpass and the type of galaxy. The rms scatter between model and Petrosian magnitudes at the bright end is now between 0.05 and 0.08 mag, depending on bandpass; the scatter between `cmodel` and Petrosian magnitudes for all galaxies is smaller, 0.03–0.05 mag. For comparison, the code that was used in the EDR and DR1 had scatters of 0.1 mag and greater, with much more significant offsets.

The `cmodel` and PSF magnitudes of stars are forced to be identical in the mean by aperture corrections; this was true in older versions of the pipeline. The rms scatter between model and PSF magnitudes for stars is much reduced, from 0.03 to 0.02 mag, with the exact values depending on bandpass. In the EDR and DR1, star-galaxy separation was based on the difference between model and PSF magnitudes (cf. the discussions by Stoughton et al. 2002, Scranton et al. 2002, and Strauss et al. 2002). We now do star-galaxy separation to set the `type` parameter in the pipeline outputs using the difference between `cmodel` and PSF magnitudes, with the threshold at the same value (0.145 mag).

Given the excellent agreement between `cmodel` magnitudes and PSF magnitudes for point sources, and between `cmodel` magnitudes and Petrosian magnitudes (albeit with intrinsic offsets due to aperture corrections) for galaxies, the `cmodel` magnitude is now an adequate proxy to use as a universal magnitude for all types of objects. Since it is approximately a matched aperture to a galaxy, it has the great advantage over Petrosian magnitudes, in particular, of having close to optimal noise properties.

For measuring *colors* of extended objects, however, we continue to recommend using the model (not the `cmodel`) magnitudes; the colors of galaxies were almost completely unaffected by the DR1 software error (cf. the discussion in § 5). The model magnitude is calculated using the best-fit parameters in the *r* band and applies it to all other bands; the

<sup>58</sup> Due to an accident of history, this parameter is misleadingly termed `fracPSF` in the flat files of the DAS.

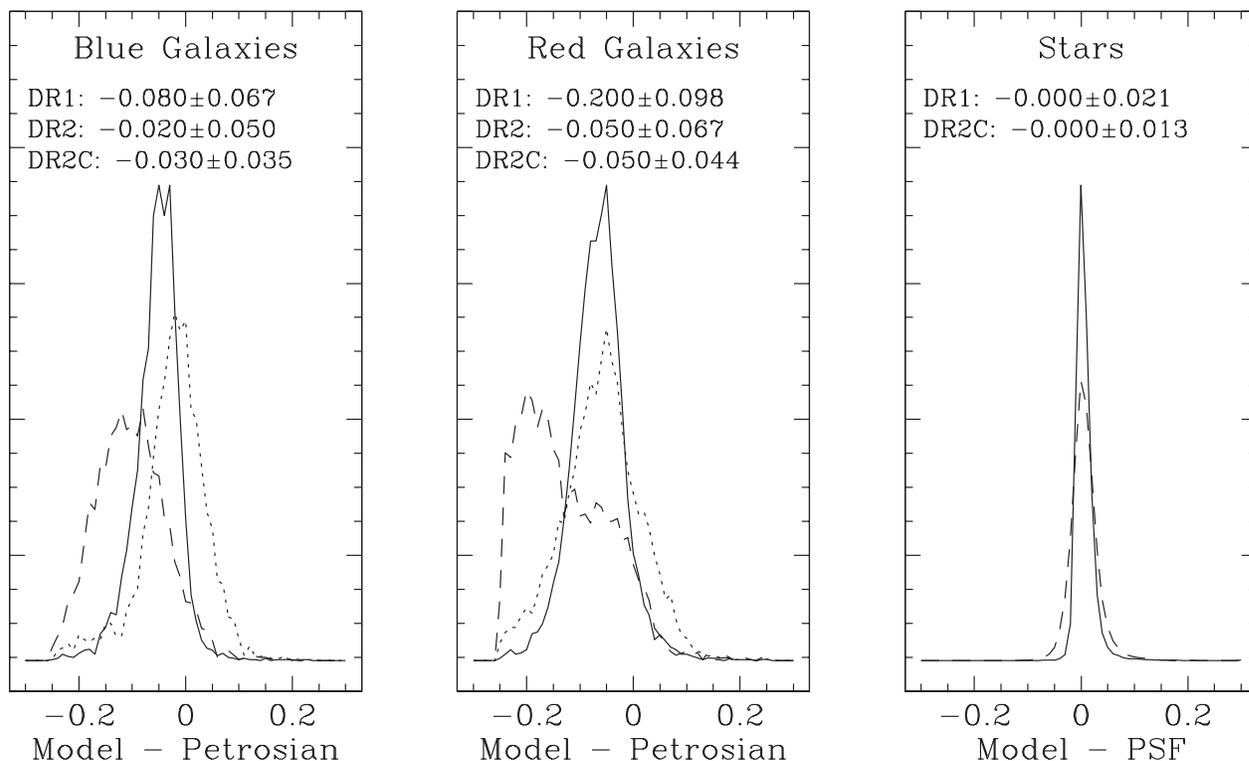


FIG. 2.—Improvements in model magnitudes for stars and galaxies in the DR2 reductions. The first panel shows the distribution of differences between  $r$  band model and Petrosian magnitudes for blue ( $u - r < 2.22$ ; Strateva et al. 2001) galaxies brighter than  $r_{\text{petro}} = 19$ ; the three curves are for the old (DR1) reductions (dashed line), the current reductions using model magnitudes (dotted line; DR2), and the current reductions using `cmodel` magnitudes (solid line; DR2C). The mode and standard deviation (based on the interquartile range) of each distribution are given. The bias in model magnitudes in the DR1 reductions is apparent. The second panel shows the same quantities for red ( $u - r > 2.22$ ) galaxies. The third panel shows the difference between `cmodel` and PSF magnitudes for  $r_{\text{PSF}} < 20$  stars, in the DR1 (dotted line) and DR2 (solid line) reductions; the width of the distribution has decreased by 40% with the new reductions.

light is therefore measured consistently through the same aperture in all bands.

### 3.2. Other Substantive Changes to the Imaging Pipelines

1. The behavior of the deblender of overlapping images has been further improved since the DR1; these changes are most important for bright galaxies of large angular extent ( $\theta > 1'$ ). In the EDR, and to a lesser extent in the DR1, bright galaxies were occasionally “shredded” by the deblender, i.e., interpreted as two or more objects and taken apart. With improvements in the code that finds the center of large galaxies in the presence of superposed stars and the deblending of stars superposed on galaxies, this shredding now rarely happens. Indeed, inspections of several hundred galaxies from the New Galaxy Catalog (NGC) shows that the deblend is correct in 95% of the cases; most of the exceptions are irregular galaxies of various sorts.

2. The PSF is measured from atlas images roughly  $7''$  across for stars; any error in the sky level determined from these images couples to spatial variability of the PSF by the Karhunen-Loève expansion used to model the PSF. This manifested itself in systematic offsets between the PSF and model magnitudes of stars of several hundredths of a magnitude, even with the fixes in the model magnitude code described above. This zero-point term in the PSF is now explicitly suppressed.

3. The pixel size is  $0''.396$ , giving well-sampled images for the typical seeing of  $1''$  or more. On rare occasions when the seeing became much better than  $0''.9$  (FWHM), the under-sampling causes the code that found stars suitable for determining the PSF to miss many objects, yielding an incorrect PSF and therefore poor stellar photometry (the seeing was

never good enough in the runs included in DR1, so this error was not triggered). Changes to the thresholds for the selection of PSF stars have solved this problem.

4. Astrometry for each object is referred to the reference frame of the  $r$ -band images. However, for objects of extreme color that are undetected in the  $r$  band (for example, cool brown dwarfs and  $z > 5.7$  quasars), DR1 had an error in the astrometric transformation from the detection band to the  $r$ -band, resulting in positional errors of several arcseconds. This problem is fixed in DR2, and the positions of objects not detected in the  $r$ -band are now correct.

5. The EDR and DR1 match each SDSS object to the nearest object in USNO-A2.0 (Monet et al. 1998), using a  $30''$  matching radius. USNO-A2.0 provides positions at a single epoch (no proper motions are provided), based on POSS-I plates. Proper motions are then calculated based on the SDSS and POSS-I positions, with a typical time baseline of 50 yr. For motions greater than  $\sim 40 \text{ mas yr}^{-1}$ , corresponding to separations between the SDSS and USNO-A2.0 positions of greater than  $2''$ , contamination by false matches becomes significant and rises with increasing motion/separation.<sup>59</sup> The DR2 reductions use USNO-B1.0 (Monet et al. 2003), which provides positions and proper motions based on various Schmidt photographic surveys (primarily POSS-I and POSS-II in the area of sky covered by SDSS). Each SDSS object is matched to the nearest USNO-B1.0 object within  $1''$ , after first converting the USNO-B1.0 positions to the epoch of the SDSS observations. This eliminates nearly all of the false matches, yielding much cleaner samples of high proper motion stars. The USNO-B1.0 proper

<sup>59</sup> See the DR1 Web site, <http://www.sdss.org/dr1>, for a fuller discussion.

motion is then given for each matching SDSS object. More sophisticated techniques, using SDSS astrometry to recalibrate the USNO-B1.0 astrometry and then recalculate the proper motions based on both SDSS and USNO-B1.0 positions, are discussed by Munn et al. (2004) and Gould & Kollmeier (2004).

6. When an image is saturated in the SDSS imaging data, the CCD wells overflow and a bleed trail results. However, the total number of electrons associated with the object, bleed trail and all, still at least approximately reflects the brightness of the object. For objects for which the flag HAS\_SATUR\_DN is set in a given band, the imaging pipeline includes the counts associated with the bleed trail of saturated objects in flux measurements. In particular, the fiber, Petrosian, PSF, and model magnitudes include this light, and it is added to the central value of the radial profile (i.e., `profMean[0]`). Since the pipeline works on a single frame at a time, bleed trails that cross frame boundaries will not be properly accounted for. In addition, the fluxes of close pairs of saturated stars whose saturated regions overlap will not be correct.

### 3.3. Newly Discovered Problems in the Imaging Data

To the best of our knowledge, none of the problems described below is so severe as to make any substantive change to the conclusions of science papers using SDSS data.

1. The  $u$  filter has a natural red leak around 7100 Å (cf. Smith et al. 2002), which is supposed to be blocked by an interference coating. However, under the vacuum in the camera, the wavelength cutoff of the interference coating has shifted redward (see the discussion by Stoughton et al. 2002), allowing some of this red leak through. The extent of this contamination is different for each camera column. It is not completely clear if the effect is deterministic; there is some evidence that it is variable from one run to another with very similar conditions in a given camera column. Roughly speaking, however, this is a 0.02 magnitude effect in the  $u$  magnitudes for mid-K stars (and galaxies of similar color), increasing to 0.06 magnitude for M0 stars ( $r - i \approx 0.5$ ), 0.2 magnitude at  $r - i \approx 1.2$ , and 0.3 magnitude at  $r - i = 1.5$ . There is a large dispersion in the red leak for the redder stars, caused by three effects:

A. The differences in the detailed red leak response from column to column, beating with the complex red spectra of these objects.

B. The almost certain time variability of the red leak.

C. The red-leak images on the  $u$  chips are out of focus and are not centered at the same place as the  $u$  image because of lateral color in the optics and differential refraction—this means that the fraction of the red-leak flux recovered by the PSF fitting depends on the amount of centroid displacement. To make matters even more complicated, this is a *detector* effect. This means that it is not the real  $u-i$  and  $u-z$  that drive the excess, but the instrumental colors (i.e., including the effects of atmospheric extinction), so the leak is worse at high air mass, when the true ultraviolet flux is heavily absorbed but the infrared flux is relatively unaffected. Given these complications, we cannot recommend a specific correction to the  $u$ -band magnitudes of red stars, and warn the user of these data about overinterpreting results on colors involving the  $u$  band for stars later than K.

2. There is a slight and only recently recognized downward bias in the determination of the sky level in the photometry, at the level of roughly 0.1 DN pixel<sup>-1</sup>. This is apparent if one compares large-aperture and PSF photometry of faint stars; the bias is of order 29 mag arcsec<sup>-2</sup> in  $r$ . This, together with

scattered light problems in the  $u$  band, can cause of order 10% errors in the  $u$ -band Petrosian fluxes of large galaxies.

3. The SDSS photometry is intended to be on the AB system (Oke & Gunn 1983), by which a magnitude 0 object should have the same counts as a flat-spectrum source of  $F_\nu = 3631$  Jy. However, this is known not to be exactly true; the SDSS photometric zero points as used in DR2 are slightly off the AB standard. We are continuing our effort to pin down these offsets. Our present estimate, based on comparison to the STIS standards of Bohlin et al. (2001) and confirmed by SDSS photometry and spectroscopy of fainter hot white dwarfs, is that the  $u$ -band zero point is in error by 0.04 mag,  $u_{AB} = u_{SDSS} - 0.04$  mag, and that  $g$ ,  $r$ , and  $i$  are close to AB. These statements are certainly not precise to better than 0.01 mag; in addition, they depend critically on the system response of the SDSS 2.5 m, which was measured by Doi et al. (2004). The  $z$ -band zero point is not as certain at this time, but there is mild evidence that it may be shifted by about 0.02 mag:  $z_{AB} = z_{SDSS} + 0.02$  mag. The large shift in the  $u$  band was expected because the adopted magnitude of the SDSS standard BD +17°4708 in Fukugita et al. (1996) was computed at zero air mass, thereby making the assumed  $u$  response bluer than that at the mean air mass. We intend to give a fuller report on the SDSS zero points, with uncertainties, in the near future. The DR2 data do *not* include the above corrections. For further discussion of the conversion between magnitudes and physical units, see the Appendix.

4. About 0.3% of the DR2 imaging footprint area (about 300 out of 100,000 fields, or 10 deg<sup>2</sup>) for DR2 are marked as “holes,” indicated in the CAS by setting quality equal to 5 (HOLE) in the “Field” table. These are areas of sky where no objects are cataloged, and researchers interested in structure statistics of galaxy or star distributions may wish to mask out these holes from their coverage map. Roughly half of these fields include a very bright star (generally  $r < 5$ ) or a very large galaxy or globular cluster, causing the object deblending in the photometric pipeline to time out. While no catalog information is available for these fields, the corrected image is available. Data of sufficiently poor quality can also be marked as a hole: very poor seeing (significantly worse than 2'' FWHM), glitches in the telescope tracking, and nonphotometric data. There are also a few small gaps, also marked as holes, which fall between two adjacent SDSS scans.

5. The  $u$  chip in the third column of the camera is read out on two amplifiers. On occasion, electronic problems on this chip caused one of the two amplifiers to fail, meaning that half the chip has no detected objects on it. This was a problem for only two of the 105 imaging runs included in DR2: run 2190, which includes a total of 360 frames in two separate contiguous pieces on strip 12N (centered roughly at  $\delta = +5^\circ$  in the north Galactic cap), and run 2189, which includes 76 frames on stripe 36N near the northern boundary of the contiguous area in the north Galactic cap. The relevant frames are flagged as bad in the `quality` flag; in addition, individual objects in this region have the  $u$  band flagged as NOTCHECKED\_CENTER (or, for objects that straddle the boundary between the two amplifiers, LOCAL\_EDGE). Richards et al. (2002) describe how the quasar selection algorithm handles such data; the net effect is that no quasars are selected by the *ugri* branch of the algorithm for these data.

## 4. THE SDSS SPECTROSCOPIC DATA

### 4.1. Improvements to Spectrophotometric Calibration

There have been three substantial improvements to the algorithms that photometrically calibrate the spectra (Tremonti

et al. 2004): (1) improved matching of observed standard stars to models; (2) tying the spectrophotometry to the  $r$ -band fiber magnitudes measured from the most recent version of the photometric pipeline; and (3) no longer using the “smear” exposures.

*Analysis of spectroscopic standard stars.*—As described by Stoughton et al. (2002), each spectroscopic plate contains 16 spectrophotometric standard stars, chosen by their colors to be F subdwarf stars. In the EDR and DR1 Spectro2d calibration pipelines, fluxing was achieved by *assuming* that the mean spectrum of the stars on each half-plate was equivalent to a synthetic composite F8 subdwarf spectrum from Pickles (1998). In the reductions included in DR2, the spectrum of each standard star is spectrally typed by comparing with a grid of theoretical spectra generated from Kurucz model atmospheres (cf. Kurucz 1992) using the spectral synthesis code SPECTRUM (Gray & Corbally 1994; Gray et al. 2001). The flux calibration vector is derived from the average ratio of each star (after correcting for Schlegel et al. [1998] reddening) and its best-fit model.

Unlike the EDR and DR1, the final calibrated DR2 spectra are *not* corrected for foreground Galactic reddening [a relatively small effect; the median  $E(B-V)$  over the survey is 0.034]. This may be changed in future data releases.

*Improved comparison to fiber magnitudes.*—We now compute the absolute calibration by tying the  $r$ -band fluxes of the standard-star spectra to the fiber magnitudes output by the latest version of the photometric pipeline. The latest version now corrects fiber magnitudes to a constant seeing of  $2''$  and includes the contribution of flux from overlapping objects in the fiber aperture; these changes greatly improve the overall data consistency.

*Smears.*—As Stoughton et al. (2002) describe, “smear” observations are low signal-to-noise ratio (S/N) spectroscopic exposures made through an effective  $5''.5 \times 9''$  aperture, aligned with the parallactic angle. Smears were designed to account for object light excluded from the  $3''$  fiber due to seeing, atmospheric refraction, and object extent. However, extensive experiments comparing photometry and spectrophotometry calibrated with and without smear observations have shown that the smear correction provides improvements only for point sources (stars and quasars) with very high S/N. For extended sources (galaxies) the spectrum obtained in the  $3''$  fiber aperture is calibrated to have the total flux and spectral shape of the light in the smear aperture. This is undesirable, for example, if the fiber samples the bulge of a galaxy, but the smear aperture includes much of its disk. For this reason, we do *not* apply the smear correction to the data in DR2.

To the extent that all point sources are centered in the fibers in the same way as are the standards, our flux calibration scheme corrects the spectra for losses due to atmospheric refraction without the use of smears. Extended sources are likely to be slightly overcorrected for atmospheric refraction. However, most galaxies are quite centrally concentrated and more closely resemble point sources than uniform extended sources. In the mean, this overcorrection makes the  $g-r$  color of the galaxy spectra too red by  $\sim 1\%$ .

Figure 3 (*left*) compares  $r$  fiber magnitudes with those synthesized from the spectra of all DR2 objects with spectral S/N per pixel greater than 5. For point sources alone, this rms difference is 0.040 mag, a 45% improvement over DR1. For extended sources, the effect of the smears was to give a systematic offset between spectroscopic and fiber magnitudes of up to a magnitude; with the DR2 reductions, this trend

is gone. The slight offset of the mean from zero is a seeing effect.

Figure 3 (*right*) compares the  $g-r$  and  $r-i$  colors derived from the spectra and photometry; the scatter is  $\sim 40\%$  lower in DR2 than in DR1. The few percent offset of the colors from zero is an indication that there are small residual errors in our spectrophotometry, perhaps due to errors in the theoretical models used to calibrate the standard stars or to offsets between our photometric system and a true AB system (see the discussion at the end of § 3).

To evaluate our spectrophotometry over smaller scales, of order  $100 \text{ \AA}$ , we compared the calibrated spectra of a sample of 166 hot DA white dwarfs drawn from Kleinman et al. (2004) with theoretical models. DA white dwarfs are useful for this comparison because they have simple hydrogen atmospheres that can be accurately modeled (e.g., Finley et al. 1997). Figure 4 shows the results of dividing each spectrum by its best-fit model. The median of the curves shows a net residual on the order of 2% at the bluest wavelengths. This is a major improvement over DR1 where the residuals were on the order of 15% at  $4300 \text{ \AA}$  because of the mismatch between the observed standard stars and the assumed model.

#### 4.2. Problematic Spectroscopic Plates

A small number of plates, listed in Table 2, suffered from a variety of minor problems. The CCD frames for several plates suffered from a transient electronic problem in the red camera in Spectrograph 2, causing the columns of the CCD to be misaligned on readout. This was fixed in software, and we believe the data to be reliable. Another set of plates labeled “Spectrograph collimation problem” suffered from having the spectrograph collimator improperly focused. This problem caused a mismatch between the flat-fields and the science exposure instrumental profile shapes on the CCD in both the spatial and wavelength directions, causing the optimal extraction process to reject an excessive number of pixels. This problem was fixed in software, and comparing overlapping objects from adjacent plates confirms that the redshifts from these problematic plates are unbiased. However, the spectra themselves should not be used for precision work or spectrophotometry. Finally, during the exposure of one plate, light from an LED somewhere on the telescope found its way to the spectrographs, resulting in an artificial excess of light centered roughly at  $6500 \text{ \AA}$ ; the spectrophotometry of this plate is quite poor.

#### 4.3. Stellar Radial Velocities

Spectra for approximately 35,000 Galactic stars of all common spectral types, targeted both by the quasar target selection algorithm, and in directed stellar programs, are available with DR2. Radial velocities (RVs) are stored as redshifts and were measured by cross-correlation to a set of stellar templates. Repeat observations of spectroscopic plates show that the stellar radial velocities are reproducible to roughly  $5 \text{ km s}^{-1}$  for stars brighter than about  $r \sim 18$ . However, the DR2 cross-correlation procedure used introduces small additional systematic errors in addition to possible dispersion errors, depending on spectral type; these systematics are of order  $10 \text{ km s}^{-1}$  or less for stars of spectral type A through K with S/N per resolution element greater than 10. For white dwarfs and low S/N A stars, lines are too broad for accurate RV determination, while M dwarf radial velocities are also less reliable due to prominent molecular bands. New

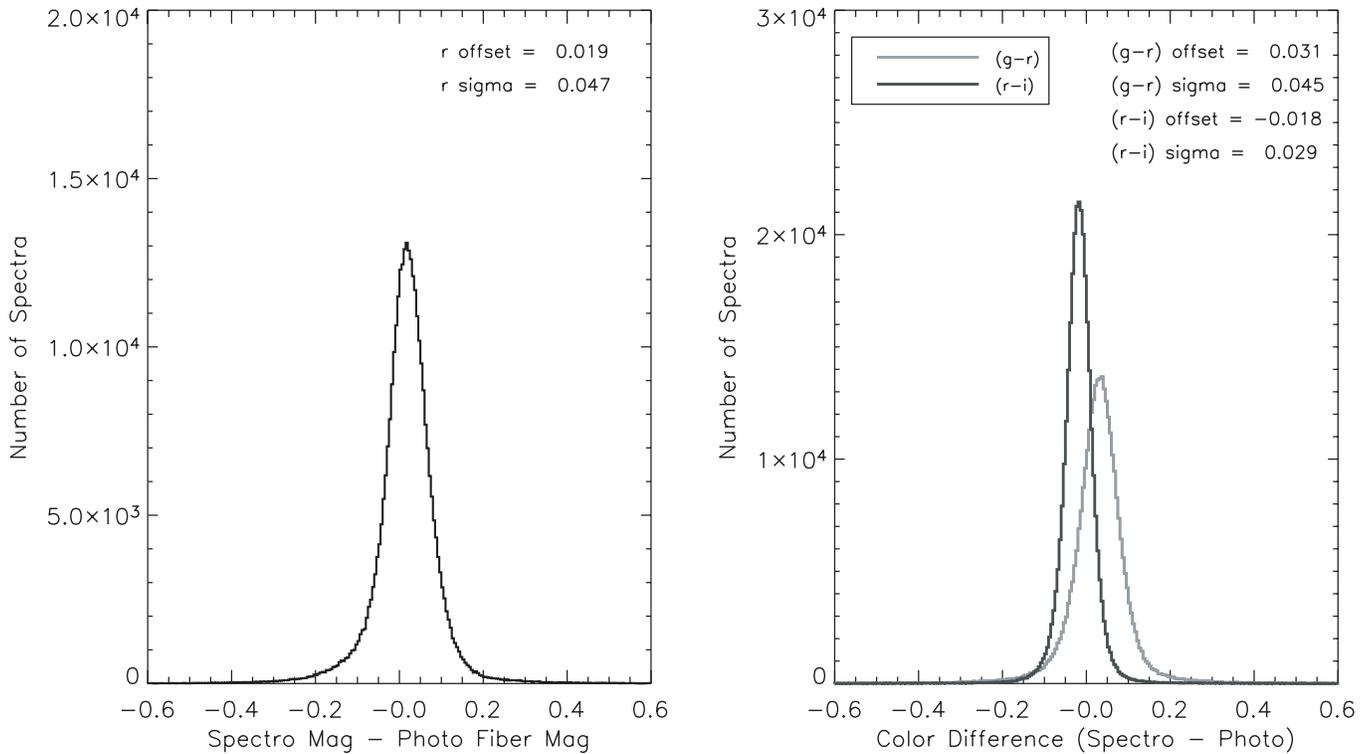


FIG. 3.—Comparison of synthetic  $r$  magnitudes and  $g-r$ ,  $r-i$  colors synthesized from the spectra with photo fiber magnitudes. We have included all objects in DR2 with S/N per pixel greater than 5.

for DR2 are measures of the centers and depths of the Ca triplet lines (8500, 8544, 8664 Å); these may be useful for more refined radial velocities of M stars.

We note that, in the EDR and DR1, zero-point errors in the templates resulted in quoted RVs for low-metallicity F stars

that were too large by  $20 \text{ km s}^{-1}$ , while the velocities of A-type stars were too large by  $49 \text{ km s}^{-1}$  (Yanny et al. 2004). These problems were uncovered through an ongoing process involving observations of known RV standards and cross checks with other model and template fitting techniques; we

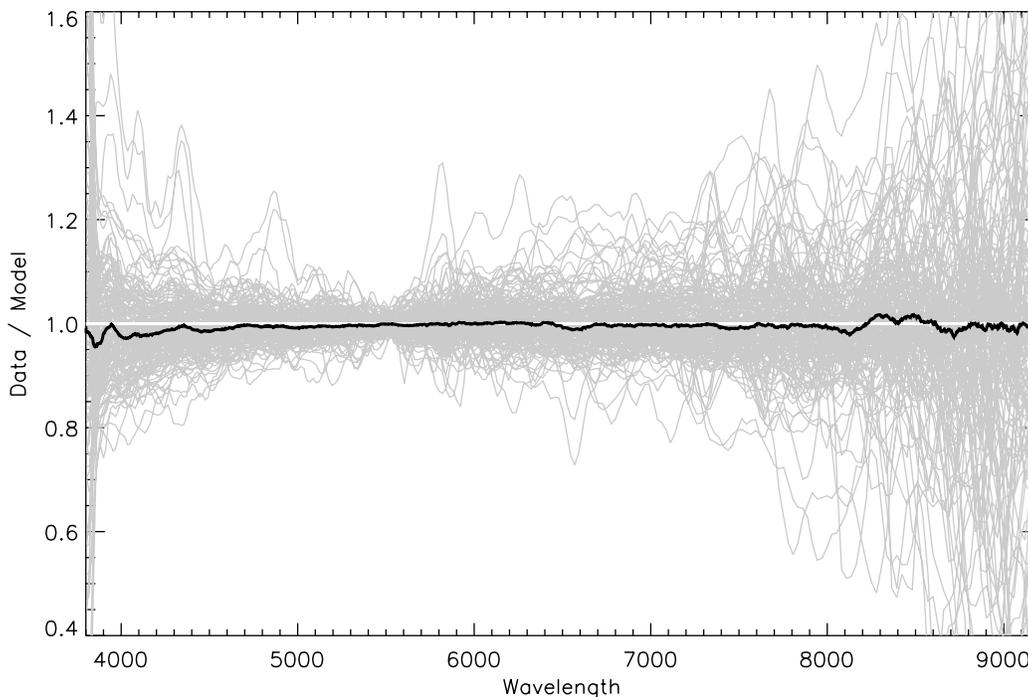


FIG. 4.—Comparison of white dwarf spectra and models. The gray lines represent 166 individual spectra divided by their best-fit model. The heavy line is the median. The equivalent median residuals in DR1 were on the order of 15% at 4300 Å; they are now on the order of a few percent.

TABLE 2  
PLATES WITH KNOWN PROBLEMS

Plate ID	Modified Julian Date	Problem
312.....	51689	Problems with readout
338.....	51694	Problems with readout
339.....	51692	Problems with readout
343.....	51692	Problems with readout
344.....	51693	Problems with readout
346.....	51693	Problems with readout
349.....	51699	Problems with readout
350.....	51691	Problems with readout
353.....	51703	Problems with readout
426.....	51882	Red light leak
504.....	52316	Spectrograph collimation problem
721.....	52228	Spectrograph collimation problem
761.....	52266	Spectrograph collimation problem
769.....	52282	Spectrograph collimation problem
770.....	52282	Spectrograph collimation problem
775.....	52295	Spectrograph collimation problem
778.....	52337	Spectrograph collimation problem

hope to continue to improve the tabulated stellar properties in future data releases or as value added data product releases.

#### 4.4. Mismatches Between the Spectra and Photometric Data

Each spectroscopic plate includes 32 sky fibers, placed in regions where the imaging data do not include a detected object; by definition, there is no photometric object associated with these. Similarly, there are 437 fibers among the 367,360 spectra in DR2 which were broken at the time of observation; neither a spectrum nor a photometric object is associated with them. Of the remainder, there are 62 objects for which the matching of fibers to objects cannot be reconstructed with any confidence, and therefore whose right ascensions and declinations are uncertain. For these objects, the right ascension and declination are listed as  $-9999$ .

Errors in the deblending algorithm in the target reductions caused spectroscopy to be carried out occasionally on non-existent objects (e.g., diffraction spikes of bright stars or satellite trails), or incorrectly shredded fragments of large galaxies (at  $z = 0.01$ , a full 10% of galaxy targets are such fragments). Many of these objects no longer exist in the best imaging reductions with its improvements to the deblender. In other cases, the photometric pipeline timed out during the best imaging reductions in fields for which target imaging proceeded without problem, so that the best photometry is missing for bona fide objects. This predominantly happens in fields close to a few very bright stars. We expect to recover objects from these “timeout holes” in future data releases.

A total of 663 spectroscopic objects therefore do not have a counterpart in the best images, 0.2% of the total. Of these, 80 (including the 62 unmapped fibers) can only be retrieved from the table `specObjA11` in the database. The remaining 583 objects are contained in the default spectroscopic table `specObj`, but will not be found by queries requesting both photometric and spectroscopic information.

#### 4.5. Redshift and Classification Correctness

As described in Abazajian et al. (2003), we have compared the results of two independent codes to measure the redshifts of the spectra. These codes are the pipeline used for the official SDSS reductions, whose redshifts are based on cross-correlation and emission-line fitting, and an independent pipe-

line that uses a  $\chi^2$  method to fit templates directly to spectra (e.g., Glazebrook et al. 1998). Only 1.7% of the objects in DR2 not included in DR1 show gross differences in redshift and/or classification between the two codes. We examined the spectra of all these cases by eye. One-third of those discrepancies are for very low S/N spectra, for which no redshift is determinable; in the vast majority of these cases, both pipelines correctly indicated that they had failed. Of the remainder, the redshifts or spectral classifications in the official reductions were clearly incorrect in 0.3% of the spectra, a few hundred objects in total (many of them intrinsically interesting objects such as extreme broad absorption line quasars, superpositions of objects, and other oddities). A list of corrections will be posted to the DR2 Web site when it is completed. A similar exercise was carried out for the DR1 data (and a similar error rate was found); the resulting corrections are incorporated into the database.

## 5. TARGET SELECTION

With the change in the model magnitude code (§ 3.1), the mean  $g-r$  and  $r-i$  model colors of galaxies have shifted by about 0.005 mag. Because the target selection for luminous red galaxies (LRGs; Eisenstein et al. 2001) is very sensitive to color, this would have increased the number density of targets by about 10%. Instead, we shifted the LRG color cuts to compensate; in addition, improved star-galaxy separation allows tighter cuts on the model-PSF quantity by which stars are rejected. Here we give the updated equations (2) and (3), respectively, from Eisenstein et al. (2001):

$$c_{\perp} = (r - i) - (g - r)/4.0 - 0.177; \quad (2)$$

$$c_{\parallel} = 0.7(g - r) + 1.2[(r - i) - 0.177]; \quad (3)$$

equations (4) and (8) for cut I:

$$r_{\text{Petro}} < 13.116 + c_{\parallel}/0.3; \quad (4)$$

$$r_{\text{PSF}} - r_{\text{model}} > 0.24; \quad (5)$$

and equations (10), (11), and (13) for cut II:

$$c_{\perp} > 0.449 - (g - r)/6; \quad (6)$$

$$g - r > 1.296 + 0.25(r - i); \quad (7)$$

$$r_{\text{PSF}} - r_{\text{model}} > 0.4. \quad (8)$$

This new version of LRG target selection is applied to the best region of sky reduced with the latest version of the imaging pipeline. It is, of course, not applied retroactively to the target version of the sky, which used older versions of the pipeline.

Due to other subtle differences in the photometric pipeline and the calibration, these changes will not exactly reproduce the selection criteria actually used when spectroscopy was carried out. Indeed, defining an LRG sample based on the best reductions will result in large spectroscopic incompleteness because so many objects are close to the boundaries. Instead, one should use the target photometry and adjust the calibrations of that relative to the best calibration. Of course, if one is interested in photometric properties of single objects, then we recommend the best photometry.

The selection of the main galaxy sample (Strauss et al. 2002) is based on Petrosian magnitudes, which have not

changed substantially with the latest versions of the pipelines. Thus, the magnitude limit for this sample,  $r_{\text{Petro}} < 17.77$ , has not changed. The improvements to the model magnitudes have allowed us to tighten the star-galaxy separation in galaxy target selection; the code uses the same cut as LRG cut I (eq. [5]). Note that some of the EDR and DR1 data were selected with other photometric limits; see the discussion in Appendix A of Tegmark et al. (2004) for details.

As described by Stoughton et al. (2002), Abazajian et al. (2003) and Schneider et al. (2003), the quasar spectroscopic target selection algorithm has evolved in the history of the SDSS. The final version described by Richards et al. (2002) went into effect just after the last of the DR1 data were taken. Thus, all DR2 data not included in DR1 (i.e., spectroscopic plate numbers greater than and including 716) use exactly the algorithm described in Richards et al. (2002). The most important change implemented at that time was the addition of sharp color cuts for high-redshift quasars.

## 6. THE FUTURE

As the name implies, DR2 is the second of a series of releases of what will eventually be the entire Sloan Digital Sky Survey. The third data release, DR3, is planned for late 2004. DR3 will include all SDSS survey-quality data taken through 2003 June, and will include on the order of a 50% increment

over DR2 in spectroscopy and imaging. We expect it to use essentially the same software used for the processing of DR2.

Funding for the creation and distribution of the SDSS Archive has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the U.S. Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS Web site is <http://www.sdss.org/>.

The SDSS is managed by the Astrophysical Research Consortium (ARC) for the Participating Institutions. The Participating Institutions are The University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, The Johns Hopkins University, Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, University of Pittsburgh, Princeton University, the US Naval Observatory, and the University of Washington. We would like to dedicate this paper to the memory of J. Beverly Oke, whose work on instrumentation and photometric calibration over the decades was crucial for the conception and development of the SDSS.

## APPENDIX

### CONVERSION BETWEEN MAGNITUDES, COUNTS, AND FLUXES

As discussed by Stoughton et al. (2002), the SDSS catalogs report asinh magnitudes (Lupton et al. 1999) instead of the conventional logarithmic magnitudes. The two magnitude definitions differ only for objects detected at low signal-to-noise ratio with the SDSS imaging camera. While asinh magnitudes produce meaningful colors for such objects, their magnitudes cannot be converted into janskys or absolute magnitudes in the same way as those of brighter objects. We therefore give detailed instructions here for the conversion between SDSS asinh magnitudes, janskys, and imaging camera counts. For reference, the definition of asinh magnitudes is

$$m = -\frac{2.5}{\ln 10} \left[ \operatorname{asinh} \left( \frac{f/f_0}{2b} \right) + \ln b \right], \quad (\text{A1})$$

where  $f_0$  is the photometric zero point of each filter and  $b$  is a softening parameter as given in Table 3 (repeated from Stoughton et al. 2002). The asinh magnitude differs by more than 1% from the conventional magnitude for objects with flux less than about  $10bf_0$ . For objects brighter than this, conversion to janskys and absolute magnitude can be done in the same way as for conventional magnitudes with less than 1% error.

To the extent that the SDSS photometry is on the AB system (§ 3.3), the photometric zero point  $f_0$  is given by 3631 Jy (Fukugita et al. 1996). With this and the assumption that the object's spectrum is flat in  $f_\nu$ , asinh magnitudes can be converted into janskys by inverting equation (1). A correction for the spectral shape can be obtained from synthetic photometry using the SDSS sensitivity curves which are available for download from the instruments/imager section of the SDSS web site.

TABLE 3  
asinh MAGNITUDE SOFTENING PARAMETERS

Band	$b$	Zero Flux Magnitude [ $m(f/f_0 = 0)$ ]	$m(f/f_0 = 10b)$
<i>u</i> .....	$1.4 \times 10^{-10}$	24.63	22.12
<i>g</i> .....	$0.9 \times 10^{-10}$	25.11	22.60
<i>r</i> .....	$1.2 \times 10^{-10}$	24.80	22.29
<i>i</i> .....	$1.8 \times 10^{-10}$	24.36	21.85
<i>z</i> .....	$7.4 \times 10^{-10}$	22.83	20.32

NOTE.—These values of the softening parameter  $b$  are set to be approximate  $1 \sigma$  of the sky noise; thus, only low-S/N measurements are affected by the difference between asinh and Pogson magnitudes. The final column gives the asinh magnitude associated with an object for which  $f/f_0 = 10b$ ; the difference between Pogson and asinh magnitudes is less than 1% for objects brighter than this.

To convert between asinh magnitudes and counts  $N$  on the SDSS imaging camera, use equation (1), replacing  $f/f_0$  as follows:

$$\frac{f}{f_0} \rightarrow \frac{N}{t_{\text{exp}}} 10^{0.4(\text{aa} + \text{kk} \times \text{airmass})} \quad (\text{A2})$$

where aa, kk, and airmass are the photometric zero point, extinction term, and airmass, respectively, for the object's field and filter from the `field` table or `tsField` file (these tables also contain the gain value), and  $t_{\text{exp}} = 53.907456$  s, the exposure time for each SDSS pixel.

#### REFERENCES

- Abazajian, K., et al. 2003, *AJ*, 126, 2018  
 Blanton, M. R., Lin, H., Lupton, R. H., Maley, F. M., Young, N., Zehavi, I., & Loveday, J. 2003, *AJ*, 125, 2276  
 Bohlin, R. C., Dickinson, M. E., & Calzetti, D. 2001, *AJ*, 122, 2118  
 Doi, M., et al. 2004, in preparation  
 Eisenstein, D. J., et al. 2001, *AJ*, 122, 2267  
 Finley, D. S., Koester, D., & Basri, G. 1997, *ApJ*, 488, 375  
 Fukugita, M., Ichikawa, T., Gunn, J. E., Doi, M., Shimasaku, K., & Schneider, D. P. 1996, *AJ*, 111, 1748  
 Glazebrook, K., Offer, A. R., & Deeley, K. 1998, *ApJ*, 492, 98  
 Gould, A., & Kollmeier, J. A. 2004, *ApJS*, 152, 103  
 Gray, R. O., & Corbally, C. J. 1994, *AJ*, 107, 742  
 Gray, R. O., Graham, P. W., & Hoyt, S. R. 2001, *AJ*, 121, 2159  
 Gunn, J. E., et al. 1998, *AJ*, 116, 3040  
 Hogg, D. W., Finkbeiner, D. P., Schlegel, D. J., & Gunn, J. E. 2001, *AJ*, 122, 2129  
 Ivezić, Ž., et al. 2002, *AJ*, 124, 2943  
 Kauffmann, G., et al. 2003, *MNRAS*, 341, 54  
 Kleinman, S., et al. 2004, *ApJ*, 607, 426  
 Kurucz, R. L. 1992, in *IAU Symp. 149, The Stellar Populations of Galaxies*, ed. B. Barbuy & A. Renzini (Dordrecht: Kluwer), 225  
 Lupton, R. H., Gunn, J. E., Ivezić, Ž., Knapp, G. R., Kent, S., & Yasuda, N. 2001, in *ASP Conf. Ser. 238, Astronomical Data Analysis Software and Systems X*, ed. F. R. Harnden, Jr., F. A. Primini, & H. E. Payne (San Francisco: ASP), 269  
 Lupton, R. H., Gunn, J. E., & Szalay, A. S. 1999, *AJ*, 118, 1406  
 Monet, D. G., et al. 1998, *USNO-A2.0 (11 CD-ROMs; Flagstaff: US Naval Obs.)*  
 ———. 2003, *AJ*, 125, 984  
 Munn, J. A., et al. 2004, *AJ*, 127, 3034  
 Newberg, H. J., et al. 2003, *ApJ*, 596, L191  
 Oke, J. B., & Gunn, J. E. 1983, *ApJ*, 266, 713  
 Pickles, A. J. 1998, *PASP*, 110, 863  
 Pier, J. R., Munn, J. A., Hindsley, R. B., Hennessy, G. S., Kent, S. M., Lupton, R. H., & Ivezić, Ž. 2003, *AJ*, 125, 1559  
 Pope, A. C., et al. 2004, *ApJ*, 607, 655  
 Richards, G. T., et al. 2002, *AJ*, 123, 2945  
 Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, 500, 525  
 Schmidt, G. D., et al. 2003, *ApJ*, 595, 1101  
 Schneider, D. P., et al. 2003, *AJ*, 126, 2579  
 Scranton, R., et al. 2002, *ApJ*, 579, 48  
 Smith, J. A., et al. 2002, *AJ*, 123, 2121  
 Stoughton, C., et al. 2002, *AJ*, 123, 485  
 Strateva, I. V., et al. 2001, *AJ*, 122, 1861  
 Strauss, M. A., et al. 2002, *AJ*, 124, 1810  
 Tegmark, M., et al. 2004, *ApJ*, 606, 702  
 Tremonti, C., et al. 2004, in preparation  
 Yanny, B., et al. 2004, *ApJ*, 605, 575  
 York, D. G., et al. 2000, *AJ*, 120, 1579  
 Zakamska, N., et al. 2003, *AJ*, 126, 2125