

## SLOAN DIGITAL SKY SURVEY: EARLY DATA RELEASE

CHRIS STOUGHTON,<sup>1</sup> ROBERT H. LUPTON,<sup>2</sup> MARIANGELA BERNARDI,<sup>3</sup> MICHAEL R. BLANTON,<sup>1,4</sup> SCOTT BURLES,<sup>1,3</sup> FRANCISCO J. CASTANDER,<sup>3,5,6,7</sup> A. J. CONNOLLY,<sup>8</sup> DANIEL J. EISENSTEIN,<sup>3,9,10,11</sup> JOSHUA A. FRIEMAN,<sup>1,3</sup> G. S. HENNESSY,<sup>12</sup> ROBERT B. HINDSLEY,<sup>13</sup> ŽELJKO IVEZIĆ,<sup>2</sup> STEPHEN KENT,<sup>1,3</sup> PETER Z. KUNSZT,<sup>14</sup> BRIAN C. LEE,<sup>1</sup> AVERY MEIKSIN,<sup>15</sup> JEFFREY A. MUNN,<sup>16</sup> HEIDI JO NEWBERG,<sup>17</sup> R. C. NICHOL,<sup>18</sup> TOM NICINSKI,<sup>1,19</sup> JEFFREY R. PIER,<sup>16</sup> GORDON T. RICHARDS,<sup>20</sup> MICHAEL W. RICHMOND,<sup>21</sup> DAVID J. SCHLEGEL,<sup>2</sup> J. ALLYN SMITH,<sup>22,23</sup> MICHAEL A. STRAUSS,<sup>2</sup> MARK SUBBARAO,<sup>3</sup> ALEXANDER S. SZALAY,<sup>14</sup> ANIRUDDHA R. THAKAR,<sup>14</sup> DOUGLAS L. TUCKER,<sup>1</sup> DANIEL E. VANDEN BERK,<sup>1</sup> BRIAN YANNY,<sup>1</sup> JENNIFER K. ADELMAN,<sup>1</sup> JOHN E. ANDERSON, JR.,<sup>1</sup> SCOTT F. ANDERSON,<sup>24</sup> JAMES ANNIS,<sup>1</sup> NETA A. BAHCALL,<sup>2</sup> J. A. BAKKEN,<sup>1</sup> MATTHIAS BARTELMANN,<sup>25</sup> STEVEN BASTIAN,<sup>1</sup> AMANDA BAUER,<sup>1,26</sup> EILEEN BERMAN,<sup>1</sup> HANS BÖHRINGER,<sup>27</sup> WILLIAM N. BOROSKI,<sup>1</sup> STEVE BRACKER,<sup>1</sup> CHARLIE BRIEGEL,<sup>1</sup> JOHN W. BRIGGS,<sup>3</sup> J. BRINKMANN,<sup>28</sup> ROBERT BRUNNER,<sup>29</sup> LARRY CAREY,<sup>24</sup> MICHAEL A. CARR,<sup>2</sup> BING CHEN,<sup>14</sup> DAMIAN CHRISTIAN,<sup>30</sup> PATRICK L. COLESTOCK,<sup>1</sup> J. H. CROCKER,<sup>14,31</sup> ISTVÁN CSABAI,<sup>14,32</sup> PAUL C. CZARAPATA,<sup>1</sup> JULIANNE DALCANTON,<sup>24</sup> ARTHUR F. DAVIDSEN,<sup>14,33</sup> JOHN ERIC DAVIS,<sup>28</sup> WALTER DEHNEN,<sup>34</sup> SCOTT DODELSON,<sup>1</sup> MAMORU DOI,<sup>35</sup> TOM DOMBECK,<sup>3</sup> MEGAN DONAHUE,<sup>30</sup> NANCY ELLMAN,<sup>5</sup> BRIAN R. ELMS,<sup>2,36</sup> MICHAEL L. EVANS,<sup>24</sup> LAURENT EYER,<sup>2</sup> XIAOHUI FAN,<sup>2,10</sup> GLENN R. FEDERWITZ,<sup>1</sup> SCOTT FRIEDMAN,<sup>14</sup> MASATAKA FUKUGITA,<sup>37</sup> ROY GAL,<sup>14</sup> BRUCE GILLESPIE,<sup>28</sup> KARL GLAZEBROOK,<sup>14</sup> JIM GRAY,<sup>38</sup> EVA K. GREBEL,<sup>34</sup> BRUCE GREENAWALT,<sup>1</sup> GRETCHEN GREENE,<sup>30</sup> JAMES E. GUNN,<sup>2</sup> ERNST DE HAAS,<sup>2</sup> ZOLTÁN HAIMAN,<sup>2,11</sup> MERLE HALDEMAN,<sup>1</sup> PATRICK B. HALL,<sup>2,39</sup> MASARU HAMABE,<sup>40</sup> BRAD HANSEN,<sup>2</sup> FREDERICK H. HARRIS,<sup>16</sup> HUGH HARRIS,<sup>16</sup> MICHAEL HARVANEK,<sup>28</sup> SUZANNE L. HAWLEY,<sup>24</sup> J. J. E. HAYES,<sup>28,41</sup> TIMOTHY M. HECKMAN,<sup>14</sup> AMINA HELMI,<sup>25</sup> ARNE HENDEN,<sup>16</sup> CRAIG J. HOGAN,<sup>24</sup> DAVID W. HOGG,<sup>4</sup> DONALD J. HOLMGREN,<sup>1</sup> JON HOLTZMAN,<sup>42</sup> CHIH-HAO HUANG,<sup>1</sup> CHARLES HULL,<sup>43</sup> SHIN-ICHI ICHIKAWA,<sup>36</sup> TAKASHI ICHIKAWA,<sup>44</sup> DAVID E. JOHNSTON,<sup>3</sup> GUINEVERE KAUFFMANN,<sup>25</sup> RITA S. J. KIM,<sup>2,14</sup> TIM KIMBALL,<sup>30</sup> E. KINNEY,<sup>28</sup> MARK KLAENE,<sup>28</sup> S. J. KLEINMAN,<sup>28</sup> ANATOLY KLYPIN,<sup>42</sup> G. R. KNAPP,<sup>2</sup> JOHN KORIENEK,<sup>1</sup> JULIAN KROLIK,<sup>14</sup> RICHARD G. KRON,<sup>1,3</sup> JUREK KRZESIŃSKI,<sup>28,45</sup> D. Q. LAMB,<sup>3</sup> R. FRENCH LEGER,<sup>1</sup> SIRILUK LIMMONGKOL,<sup>24</sup> CARL LINDENMEYER,<sup>1</sup> DANIEL C. LONG,<sup>28</sup> CRAIG LOOMIS,<sup>28</sup> JON LOVEDAY,<sup>46</sup> BRYAN MACKINNON,<sup>1,47</sup> EDWARD J. MANNERY,<sup>24</sup> P. M. MANTSCH,<sup>1</sup> BRUCE MARGON,<sup>24,30</sup> PEREGRINE MCGEHEE,<sup>48</sup> TIMOTHY A. MCKAY,<sup>22</sup> BRIAN McLEAN,<sup>30</sup> KRISTEN MENOU,<sup>2</sup> ARONNE MERELLI,<sup>18</sup> H. J. MO,<sup>25</sup> DAVID G. MONET,<sup>16</sup> OSAMU NAKAMURA,<sup>37</sup> VIJAY K. NARAYANAN,<sup>2</sup> THOMAS NASH,<sup>1</sup> ERIC H. NEILSEN, JR.,<sup>1</sup> PETER R. NEWMAN,<sup>28</sup> ATSUKO NITTA,<sup>28</sup> MICHAEL ODENKIRCHEN,<sup>34</sup> NORIO OKADA,<sup>36</sup> SADANORI OKAMURA,<sup>49</sup> JEREMIAH P. OSTRIKER,<sup>2</sup> RUSSELL OWEN,<sup>24</sup> A. GEORGE PAULS,<sup>2</sup> JOHN PEOPLES,<sup>1</sup> R. S. PETERSON,<sup>1</sup> DONALD PETRAVICK,<sup>1</sup> ADRIAN POPE,<sup>14,18</sup> RUTH PORDES,<sup>1</sup> MARC POSTMAN,<sup>30</sup> ANGELA PROSAPIO,<sup>1</sup> THOMAS R. QUINN,<sup>24</sup> RON RECHENMACHER,<sup>1</sup> CLAUDIO H. RIVETTA,<sup>1</sup> HANS-WALTER RIX,<sup>34</sup> CONSTANCE M. ROCKOSI,<sup>3</sup> ROBERT ROSNER,<sup>3</sup> KURT RUTHMANSDORFER,<sup>1</sup> DALE SANDFORD,<sup>50</sup> DONALD P. SCHNEIDER,<sup>20</sup> RYAN SCRANTON,<sup>3</sup> MAKI SEKIGUCHI,<sup>37</sup> GARY SERGEY,<sup>1</sup> RAVI SHETH,<sup>1</sup> KAZUHIRO SHIMASAKU,<sup>49</sup> STEPHEN SMEE,<sup>14,51</sup> STEPHANIE A. SNEDDEN,<sup>28</sup> ALBERT STEBBINS,<sup>1</sup> CHRISTOPHER STUBBS,<sup>24</sup> ISTVÁN SZAPUDI,<sup>52</sup> PAULA SZKODY,<sup>24</sup> GYULA P. SZOKOLY,<sup>14</sup> SERGE TABACHNIK,<sup>2</sup> ZLATAN TSVETANOV,<sup>14</sup> ALAN UOMOTO,<sup>14</sup> MICHAEL S. VOGELEY,<sup>53</sup> WOLFGANG VOGES,<sup>27</sup> PATRICK WADDELL,<sup>24</sup> RENÉ WALTERBOS,<sup>42</sup> SHU-I WANG,<sup>3</sup> MASARU WATANABE,<sup>54</sup> DAVID H. WEINBERG,<sup>55</sup> RICHARD L. WHITE,<sup>30</sup> SIMON D. M. WHITE,<sup>25</sup> BRIAN WILHITE,<sup>1,3</sup> DAVID WOLFE,<sup>30</sup> NAOKI YASUDA,<sup>36</sup> DONALD G. YORK,<sup>3,56</sup> IDIT ZEHAVI,<sup>1</sup> AND WEI ZHENG<sup>14</sup>

Received 2001 September 27; accepted 2001 October 4

<sup>1</sup> Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, IL 60510.

<sup>2</sup> Princeton University Observatory, Peyton Hall, Princeton, NJ 08544-1001.

<sup>3</sup> Department of Astronomy and Astrophysics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637.

<sup>4</sup> Department of Physics, New York University, 4 Washington Place, New York, NY 10003.

<sup>5</sup> Department of Physics, Yale University, P.O. Box 208121, New Haven, CT 06520.

<sup>6</sup> Departamento de Astronomía, Universidad de Chile, Casilla 36-D, Santiago, Chile.

<sup>7</sup> Andes Prize Fellow.

<sup>8</sup> Department of Physics and Astronomy, University of Pittsburgh, 3941 O'Hara Street, Pittsburgh, PA 15260.

<sup>9</sup> Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721.

<sup>10</sup> Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540.

<sup>11</sup> Hubble Fellow.

<sup>12</sup> US Naval Observatory, 3450 Massachusetts Avenue, NW, Washington, DC 20392-5420.

<sup>13</sup> Remote Sensing Division, Code 7215, Naval Research Laboratory, 4555 Overlook Avenue, SW, Washington, DC 20375.

<sup>14</sup> Department of Physics and Astronomy, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218-2686.

<sup>15</sup> Institute for Astronomy, Royal Observatory, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK.

<sup>16</sup> US Naval Observatory, Flagstaff Station, P.O. Box 1149, Flagstaff, AZ 86002.

<sup>17</sup> Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, 110 Eighth Street, Troy, NY 12180-3590.

<sup>18</sup> Department of Physics, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15232.

<sup>19</sup> Lucent Technologies, 2000 North Naperville Road, Naperville, IL 60566.

<sup>20</sup> Department of Astronomy and Astrophysics, 525 Davey Laboratory, Pennsylvania State University, University Park, PA 16802.

<sup>21</sup> Department of Physics, Rochester Institute of Technology, 85 Lomb Memorial Drive, Rochester, NY 14623-5603.

<sup>22</sup> Department of Physics, University of Michigan, 500 East University, Ann Arbor, MI 48109-1120.

<sup>23</sup> Department of Physics and Astronomy, University of Wyoming, P.O. Box 3905, Laramie, WY 82071.

<sup>24</sup> Department of Astronomy, University of Washington, Box 351580, Seattle, WA 98195.

<sup>25</sup> Max-Planck-Institut für Astrophysik, Postfach 1317, D-85741 Garching, Germany.

## ABSTRACT

The Sloan Digital Sky Survey (SDSS) is an imaging and spectroscopic survey that will eventually cover approximately one-quarter of the celestial sphere and collect spectra of  $\approx 10^6$  galaxies, 100,000 quasars, 30,000 stars, and 30,000 serendipity targets. In 2001 June, the SDSS released to the general astronomical community its early data release, roughly 462 deg<sup>2</sup> of imaging data including almost 14 million detected objects and 54,008 follow-up spectra. The imaging data were collected in drift-scan mode in five bandpasses (*u*, *g*, *r*, *i*, and *z*); our 95% completeness limits for stars are 22.0, 22.2, 22.2, 21.3, and 20.5, respectively. The photometric calibration is reproducible to 5%, 3%, 3%, 3%, and 5%, respectively. The spectra are flux- and wavelength-calibrated, with 4096 pixels from 3800 to 9200 Å at *R* ≈ 1800. We present the means by which these data are distributed to the astronomical community, descriptions of the hardware used to obtain the data, the software used for processing the data, the measured quantities for each observed object, and an overview of the properties of this data set.

*Key words:* atlases — catalogs — surveys

## 1. INTRODUCTION

In 1988, a team of astrophysicists gathered together for the task of designing a next-generation redshift survey—one that would target both galaxies and quasars. In order to achieve the highest level of homogeneity in these two redshift samples, it was concluded that a dedicated imaging survey would be needed from which target galaxies and quasars would be selected, and that imaging and spectroscopy could be done with the same telescope switching between the two observing modes. Substantial improvement beyond existing surveys dictated an increase by a factor of 100 in terms of the number of targets available at the time—in other words, a survey of 1 million galaxy redshifts. This survey, the Sloan Digital Sky Survey (SDSS; York et al. 2000), is now underway, having begun standard operations in 2000 April, and is planned to last 5 years. It will eventually cover  $\pi$  steradians in the north Galactic cap, plus

three smaller regions in the south Galactic cap. Now, at the end of the SDSS's first year of standard operations, we are pleased to present this early data release (EDR), consisting of 462 square degrees of imaging data and 54,008 spectra of objects selected from within this area.

This is the first substantive public release of data from the SDSS. Release of the future survey data is scheduled to follow this first release in approximately annual installments. The EDR is served over the World Wide Web from the Space Telescope Science Institute,<sup>57</sup> Fermilab,<sup>58</sup> the National Astronomical Observatory of Japan,<sup>59</sup> and the Max-Planck-Institut für Astrophysik.<sup>60</sup> The institutions involved in the survey and the survey funding sources may be found at the end of this paper.

A brief description of the hardware and associated software may be found in York et al. (2000), which is a technical summary of the project and serves as an introduction to the SDSS Project Book, which is a full technical description of

<sup>26</sup> Department of Physics, University of Cincinnati, P.O. Box 210011, Cincinnati, OH 45221.

<sup>27</sup> Max-Planck-Institut für extraterrestrische Physik, Postfach 1312, D-85741 Garching, Germany.

<sup>28</sup> Apache Point Observatory, P.O. Box 59, Sunspot, NM 88349.

<sup>29</sup> Department of Astronomy, 105-24, California Institute of Technology, 1201 East California Boulevard, Pasadena, CA 91125.

<sup>30</sup> Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218.

<sup>31</sup> Ball Aerospace and Technologies Corporation, P.O. Box 1062, Boulder, CO 80301.

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

<sup>33</sup> Deceased.

<sup>34</sup> Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany.

<sup>35</sup> Institute of Astronomy and Research Center for the Early Universe, University of Tokyo, 2-21-1 Osawa, Mitaka, Tokyo 181-0015, Japan.

<sup>36</sup> National Astronomical Observatory, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan.

<sup>37</sup> Institute for Cosmic Ray Research, University of Tokyo, 5-1-5 Kashiwa, Kashiwa City, Chiba 277-8582, Japan.

<sup>38</sup> Microsoft Research, 301 Howard Street, No. 830, San Francisco, CA 94105.

<sup>39</sup> Departamento de Astronomía y Astrofísica, Pontificia Universidad Católica de Chile, Casilla 306, Santiago 22, Chile.

<sup>40</sup> Department of Mathematical and Physical Sciences, Japan Women's University, 2-8-1 Mejirodai, Bunkyo, Tokyo 112-8681, Japan.

<sup>41</sup> Institute of Astrophysics and Computational Sciences, Department of Physics, Catholic University of America, Washington, DC 20064.

<sup>42</sup> Department of Astronomy, MSC 4500, New Mexico State University, P.O. Box 30001, Las Cruces, NM 88003.

<sup>43</sup> Observatories of the Carnegie Institution of Washington, 813 Santa Barbara Street, Pasadena, CA 91101.

<sup>44</sup> Astronomical Institute, Tohoku University, Aramaki, Aoba, Sendai 980-8578, Japan.

<sup>45</sup> Obserwatorium Astronomiczne na Suhorze, Akademia Pedagogiczna w Krakowie, ulica Podchorążych 2, PL-30-084 Kraków, Poland.

<sup>46</sup> Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QJ, UK.

<sup>47</sup> Merrill Lynch Japan Incorporated, 1-1-3 Otemachi, Chiyoda, Tokyo 100, Japan.

<sup>48</sup> Los Alamos National Laboratory, P.O. Box 1663, Los Alamos, NM 87545.

<sup>49</sup> Department of Astronomy and Research Center for the Early Universe, University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan.

<sup>50</sup> Yerkes Observatory, University of Chicago, 373 West Geneva Street, Williams Bay, WI 53191.

<sup>51</sup> Department of Astronomy, University of Maryland, College Park, MD 20742-2421.

<sup>52</sup> Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822.

<sup>53</sup> Department of Physics, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104.

<sup>54</sup> Institute of Space and Astronautical Science, 3-1-1 Yoshinodai, Sagamihara, Kanagawa 229-8510, Japan.

<sup>55</sup> Department of Astronomy, Ohio State University, 140 West 18th Avenue, Columbus, OH 43210-1173.

<sup>56</sup> Enrico Fermi Institute, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637.

<sup>57</sup> See <http://archive.stsci.edu/sdss>.

<sup>58</sup> See <http://www.sdss.org>.

<sup>59</sup> See <http://sdss.nao.ac.jp>.

<sup>60</sup> See <http://www.sdss.mpg.de>.

TABLE 1  
SDSS EQUIPMENT SUMMARY

Parameter	Value
Telescope and Site: Apache Point Observatory	
Latitude and longitude .....	N32°46'49".3, W105°49'13".5
Elevation .....	2788 m
Survey telescope .....	2.5 m diameter, modified Ritchey-Chrétien design; 27% central obscuration
Survey area .....	North Galactic cap, 10,000 deg <sup>2</sup> , minimal Galactic extinction, plus three stripes in south Galactic cap
Instruments.....	Imaging camera and two double spectrographs
Photometric telescope .....	20 inch (0.5 m), with one CCD camera, filter wheel, and shutter
Imaging Camera	
Photometric CCDs .....	30, 2048 × 2048, SITe/Tektronix, 49.2 mm square
CCD read noise.....	<5 e <sup>-</sup> pixel <sup>-1</sup> (overall system is sky limited)
Image frame size .....	2048 × 1361 pixels (13'52 × 8'98)
Image column separation .....	25'17
Detector separation along column ...	17'98
Focal-plane image scale .....	3.616 mm arcmin <sup>-1</sup>
Detector image scale .....	3.636 mm arcmin <sup>-1</sup>
Pixel size and scale.....	24 μm; 0"396 pixel <sup>-1</sup>
Filters .....	<i>riuzg</i> scanned in that order, 71.7 s apart
Integration time .....	54 s
Operating mode .....	Time-delay and integrate ("drift scan")
Field distortion .....	<0".1 over entire field
Field size .....	2°5
Flux calibration .....	Standard-star fields at 15° intervals along scans, tied to BD +17°4708, atmospheric extinction determined by PT
Astrometric CCDs .....	22, 0.25 × 2 inches, above and below CCD columns; <i>r</i> filter plus 3 mag neutral density filter, 10.5 s integration time
Spectrographs	
Channels.....	One red, one blue for each spectrograph
CCDs.....	SITe/Tektronix (as for imager)
Coverage.....	3800–6150 Å (blue), 5800–9200 Å (red), $\lambda/\Delta\lambda \approx 1800$
Number of fibers .....	320 × 2
Fiber diameter .....	3"
Flux calibration .....	Standard stars in each field, tied to colors observed with camera
Integration time .....	45 minutes, in three exposures [nominal $(S/N)^2 > 15$ pixel <sup>-1</sup> at $g^* = 20.2$ ]
Pixel size .....	69 km s <sup>-1</sup>
Wavelength calibration .....	Hg, Cd, and Ne arc lamps, rms error of 0.07 pixels (10 km s <sup>-1</sup> )
Flat field.....	Quartz lamps

the survey hardware and software, available on the Web.<sup>61</sup> The as-built instrument parameters are given in Table 1, and basic characteristics of the data are given in Table 2. In brief, the survey uses a dedicated 2.5 m telescope, located at Apache Point Observatory (APO) in New Mexico, with a 3° field of view. The telescope has two instruments: a CCD imaging camera that takes data in drift-scanning mode, nearly simultaneously in five photometric bands, *u*, *g*, *r*, *i*, and *z*, and a pair of double spectrographs that use fiber optics to simultaneously take spectra of 640 objects selected from the imaging data. The imaging data are taken on nights of pristine conditions (photometric, good seeing, no Moon), while spectroscopy is done on those nights that are less than perfect. The data are photometrically calibrated with the aid of an auxiliary 20 inch (0.5 m) telescope, the photometric telescope (PT), at the site. The data are processed through a series of interlocking pipelines that find the objects in the imaging data, measure their properties, apply astrometric and photometric calibrations, select objects for

spectroscopic follow-up, extract and calibrate the spectra, and derive redshifts and spectral types from the spectra.

The data included in the EDR were taken as we commissioned the hardware and software of the survey and do not all meet our scientific requirements, in particular in image quality ("seeing"), photometric calibration, and target selection. Nevertheless, the data are of excellent quality and have supported a number of investigations: the discovery of high-redshift quasars (Anderson et al. 2001 and references therein); the large-scale distribution of galaxies (Zehavi et al. 2001); the gravitational lensing masses of galaxies (Fischer et al. 2000) and clusters (Castander et al. 2001; Sheldon et al. 2001); the luminosities and colors of galaxies (Blanton et al. 2001; Shimasaku et al. 2001; Strateva et al. 2001); the structure of the Milky Way (Ivezic et al. 2000; Yanny et al. 2000; Chen et al. 2001); the discovery of brown dwarfs (Leggett et al. 2000 and references therein); the structure of the asteroid belt (Ivezic et al. 2001); and many other results as well.<sup>62</sup> Our aim in this paper is to describe the data in enough detail to allow the community to reproduce the

<sup>61</sup> See <http://www.astro.princeton.edu/PBOOK/welcome.htm>.

<sup>62</sup> An up-to-date, complete list of SDSS science publications may be found at <http://www.sdss.org/science/pubs.html>.

TABLE 2  
SDSS EDR DATA-QUALITY SUMMARY

Parameters	Performance
Imaging	
PSF (FWHM).....	1.2''–2.3''
95% completeness limit for point source .....	$u^*, 22.0; g^*, 22.2; r^*, 22.2; i^*, 21.3; z^*, 20.5$
Sky coverage within survey area .....	98% (omitting areas around bright stars)
Photometric calibration errors for point sources .....	$u^*, 5\%; g^*, 3\%; r^*, 3\%; i^*, 3\%; z^*, 5\%$
Star-galaxy separation .....	95% correct at magnitude 21 in $r^*$
Astrometric errors.....	Absolute, <100 mas; relative, <50 mas (both rms per coordinate)
Spectroscopy	
Limiting magnitude .....	$S/N = 4.5 \text{ pixel}^{-1}$ at $\lambda f_\lambda = 10^{-13} \text{ ergs cm}^{-2} \text{s}^{-1}$
Error in galaxy redshifts.....	$\Delta z = 0.0001 (\approx 30 \text{ km s}^{-1})$
Error in quasar redshifts .....	$\Delta z = 0.001$
Success rate, redshifts for galaxies .....	>99%
Success rate, redshifts for quasars .....	98%

results of these papers and carry out further investigations with them.

The outline of this paper is as follows: Section 2 describes the scope of the EDR, the basic data formats, and the way in which the data will be distributed to the astronomical community. Section 3 describes the hardware of the project, emphasizing those characteristics that are necessary to understand the strengths and flaws of the data. Section 4 describes the pipelines used to reduce the data, with emphasis on the nature of the scientifically useful outputs. We conclude in § 5.

Finally, a comment on notation: As we describe in § 4.5, our photometric calibration remains uncertain, in part because of differing filter curves on the 2.5 m and the PT. The original filter system, and the AB system based on it as defined by Fukugita et al. (1996), is referred to here as  $u'g'r'i'z'$ ; this is close to that realized on the PT. The 2.5 m filters themselves are referred to as  $ugriz$ , while the still-preliminary 2.5 m based photometry will be called  $u^*g^*r^*i^*z^*$ .

## 2. DATA DISTRIBUTION

### 2.1. Sky Coverage

The EDR contains  $462 \text{ deg}^2$  of imaging data in five bands and 54,008 spectra in that same area. The data were acquired in three regions: along the celestial equator in the southern Galactic sky; along the celestial equator in the northern Galactic sky; and in a region overlapping the *SIRTF* First Look Survey.

Table 3 summarizes the imaging data included in the EDR. The “run number” is a designation we use for one continuous scan of the SDSS imaging camera on the sky, and a “stripe” is the great circle covered by two runs,  $2.5^\circ$  wide. We cover each “stripe” in two “strips,” separated in the north-south direction so that the interleaved scans of the six columns of the imaging camera completely cover the “stripe.” We define the great circles for the imaging survey in § 3.2.2. The location of each run and the effective area covered are indicated in the table. Runs 94/125 and 752/

TABLE 3  
EARLY DATA RELEASE IMAGING DATA

Run	Strip	Net Area (deg $^2$ )	Location (J2000) (deg)	$\lambda$ (deg)	$\eta$ (deg)	Start Field	End Field
1998 September: South Equatorial Stripe							
94.....	82N	83	$\alpha = 351\text{--}56, \delta = 0$	[166, -129] <sup>a</sup>	-32.5	104	544
125.....	82S	83				11	451
1999 March: North Equatorial Stripe							
752.....	10S	114	$\alpha = 145\text{--}236, \delta = 0$	[-40, 51]	-32.5	11	617
756.....	10N	114				196	802
2000 April: <i>SIRTF</i> First Look Survey							
1336 ...	42N	16	$\alpha = 257, \delta = 59^b$	[23, 36]	47.5	11	95
1339 ...	42S	16				11	95
1356 ...	43N	17	$\alpha = 261, \delta = 60$	[22.8, 35.8]	50	20	113
1359 ...	43S	19				19	113
Total .....		462					

<sup>a</sup> The  $\lambda$ -coordinate wraps at  $180^\circ \equiv -180^\circ$ .

<sup>b</sup> Center of the First Look area.

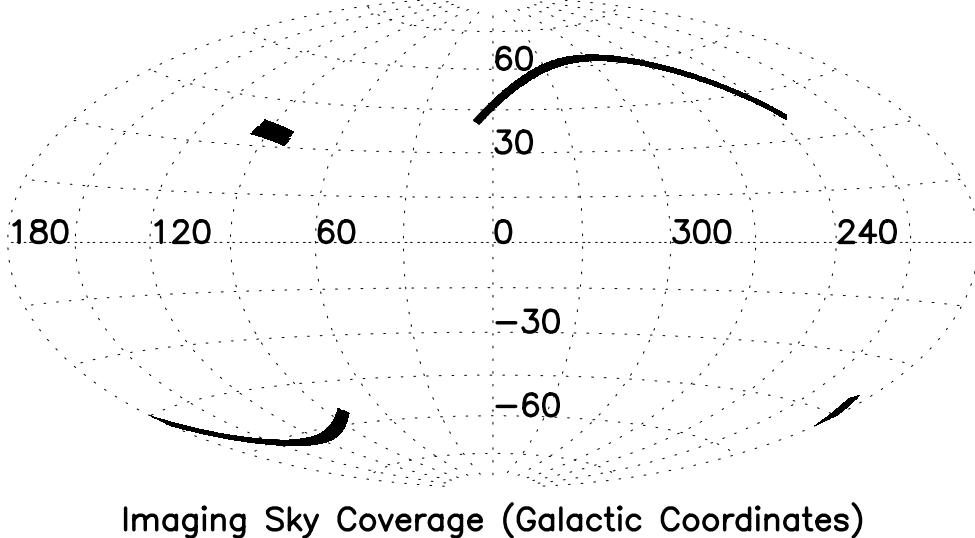


FIG. 1.—Imaging sky coverage for the SDSS EDR. A box defined by  $\text{raMin}$ ,  $\text{raMax}$ ,  $\text{decMin}$ , and  $\text{decMax}$  is filled in for each *Field* object.

756 are long stripes on the equator, in the southern and northern Galactic caps, respectively, while runs 1336/1339 and 1356/1359 are shorter scans, off the equator, designed to overlap with the *SIRTF* First Look Survey. The sky coverage of the resulting imaging and spectroscopic data is illustrated in Figures 1 and 2.

Table 4 summarizes the spectroscopic data. As discussed in § 4.8, we select objects detected in the imaging data for spectroscopic observations. The nominal exposure time for each plate is 45 minutes, which typically yields a signal-to-noise ratio of  $4.5 \text{ pixel}^{-1}$  for objects with a  $g^*$  magnitude of 20.2. The measured signal-to-noise ratio at  $g^* = 20.2$  for each plate is included in Table 4. For completeness, we list several plates that were designed but not observed; they have no entry in the  $(\text{S}/\text{N})^2$  column.

The overall quality of the EDR is summarized briefly in Table 2.

## 2.2. Data Products

Figure 3 summarizes the data products for imaging and spectroscopy, the three database servers we use to present these products to the astronomical community, and the user interfaces we have developed to help astronomers work with the data effectively. The data products described below are

*Image parameters:* The positions, fluxes, and shapes of all detected objects;

*Spectroscopic parameters:* The redshift, spectral classification, and detected lines of each spectrum;

*Color images:* JPEG images constructed from the  $g$ ,  $r$ , and  $i$  imaging data;

*Images:* FITS image files of the corrected frames (§ 3.5) in five bands, a mask that records how each pixel was used in the imaging pipelines,  $4 \times 4$  binned images of the corrected frames after detected objects have been

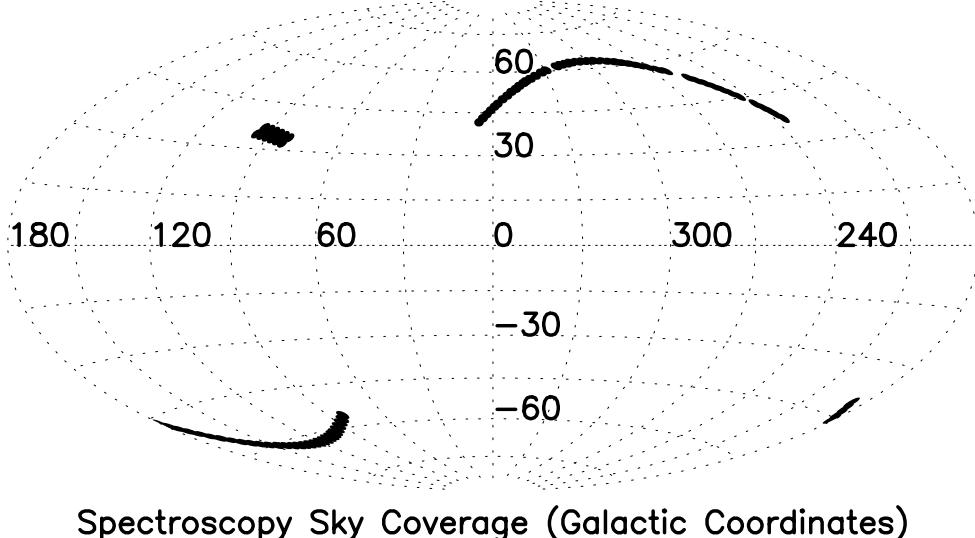


FIG. 2.—Spectroscopy sky coverage for the SDSS EDR. A circle defined by  $\text{ra}$ ,  $\text{dec}$ , and a diameter of  $3^\circ$  is filled in for each *Plate* object.

TABLE 4  
EARLY DATA RELEASE SPECTROSCOPIC PLATES

Plate ID	Tile ID	Date Completed (MJD)	$\alpha$ (J2000) (deg)	$\delta$ (J2000) (deg)	(S/N) <sup>2</sup> at $g^* = 20.2$ (pixel <sup>-1</sup> )	Smear Exposure Included?
266.....	122	51,630	145.8928	0.0613	33.8	No
267.....	121	51,608	147.7332	-0.0307	31.1	No
268.....	120	51,633	149.0618	0.0797	33.1	No
269.....	119	51,910	150.6295	0.0000	22.8	Yes
270.....	118	51,909	152.4607	0.0061	29.4	No
271.....	117	...	154.2643	0.0245	...	...
272.....	116	51,941	156.0189	0.0184	24.8	No
273.....	115	51,957	157.9021	0.0092	17.9	No
274.....	114	51,913	159.7487	0.0184	65.7	Yes
275.....	113	51,910	161.4664	-0.0061	37.3	Yes
276.....	112	51,909	163.3804	0.0429	47.1	Yes
277.....	111	51,908	165.2300	0.0061	28.3	Yes
278.....	110	51,900	167.0275	0.0276	47.1	Yes
279.....	109	...	168.8618	-0.0061	...	...
280.....	108	...	170.4477	0.1655	...	...
281.....	107	51,614	171.8773	0.1104	40.8	No
282.....	106	51,658	173.7698	-0.1196	33.8	No
283.....	105	51,959	175.9750	-0.0031	28.6	No
284.....	104	51,943	177.9810	-0.0644	27.8	No
285.....	103	51,930	179.5608	-0.0153	44.0	Yes
286.....	102	51,999	181.3061	0.0092	16.5	No
287.....	101	52,023	183.1557	0.0184	24.9	No
288.....	100	52,000	184.8459	-0.0184	20.5	No
289.....	99	51,990	186.7078	-0.0184	22.7	No
290.....	98	51,941	188.9313	0.0307	20.4	No
291.....	97	51,928	190.7257	-0.0368	31.8	Yes
292.....	96	51,609	192.5814	-0.0245	21.4	No
293.....	95	51,689	194.4555	0.0031	22.1	No
294.....	94	51,986	196.5075	0.0889	22.8	No
295.....	93	51,985	198.0996	-0.0184	17.3	No
296.....	92	51,665	199.8694	0.0674	17.3	No
297.....	91	51,959	201.5658	-0.0061	22.0	Yes
298.....	90	...	203.4644	-0.0061	...	...
299.....	89	51,671	205.3569	0.0061	31.9	No
300.....	88	51,666	207.2249	-0.0123	23.4	No
301.....	87	51,942	209.6201	-0.0552	30.9	Yes
302.....	86	51,688	211.7208	-0.0490	22.9	No
303.....	85	51,615	213.4171	0.0674	24.2	No
304.....	84	51,609	214.5129	0.0000	17.8	Yes
305.....	83	51,613	216.3073	0.0337	24.2	No
306.....	82	51,637	217.8810	-0.0828	30.1	No
307.....	81	51,663	219.7244	0.0552	24.9	No
308.....	80	51,662	221.6415	-0.0613	28.2	No
309.....	79	51,994	223.6321	0.0245	32.0	No
310.....	78	51,990	225.8618	0.0061	18.1	No
311.....	77	51,665	227.2822	0.0245	22.5	No
312.....	76	51,689	229.0766	0.0245	20.5	No
313.....	75	51,673	230.9047	-0.0215	22.7	No
314.....	74	51,641	232.7451	0.0123	16.3	No
315.....	73	51,663	234.5733	-0.0031	18.6	No
324.....	97	51,666	190.7257	-0.0368	20.6	No
349.....	148	51,699	254.1442	63.6530	47.4	No
350.....	149	51,691	258.4975	65.1336	21.8	No
351.....	150	51,780	255.9522	61.6193	22.0	No
352.....	151	51,694	260.0012	63.0373	19.9	No
353.....	152	51,703	257.5362	59.5647	37.1	No
354.....	153	51,792	261.3024	60.9272	15.4	Yes
355.....	154	51,788	258.9371	57.4932	28.9	No
356.....	155	...	262.4414	58.8061	...	...
357.....	156	...	260.1872	55.4078	...	...
358.....	157	51,818	263.4490	56.6762	26.5	Yes
359.....	158	51,821	261.3115	53.3110	18.9	Yes
360.....	159	51,816	264.3484	54.5390	21.6	Yes

TABLE 4—Continued

Plate ID	Tile ID	Date Completed (MJD)	$\alpha$ (J2000) (deg)	$\delta$ (J2000) (deg)	(S/N) <sup>2</sup> at $g^* = 20.2$ (pixel <sup>-1</sup> )	Smear Exposure Included?
362.....	159	51,999	264.3484	54.5390	27.9	No
363.....	74	51,989	232.7451	0.0123	30.1	No
366.....	155	52,017	262.4414	58.8061	18.4	No
367.....	156	51,997	260.1872	55.4078	51.2	No
382.....	171	...	349.2559	0.0289	...	...
383.....	172	51,818	351.2021	0.1141	28.3	Yes
384.....	173	51,821	353.2949	-0.0398	25.6	Yes
385.....	174	51,877	355.3896	0.0444	52.4	No
386.....	175	51,788	357.6322	0.0540	16.6	No
387.....	176	51,791	359.8243	0.0243	35.3	Yes
388.....	177	51,793	1.8246	-0.0050	18.4	Yes
389.....	178	51,795	3.5598	-0.0248	26.3	No
390.....	179	51,900	5.2229	-0.0332	53.3	Yes
391.....	180	51,782	6.9964	0.0453	18.7	No
392.....	181	51,793	8.8956	-0.0021	26.9	Yes
393.....	182	51,794	10.8188	0.0369	23.6	Yes
394.....	183	51,876	12.7276	-0.0330	38.7	Yes
395.....	184	51,783	14.6091	0.0160	22.3	No
396.....	185	51,816	16.4703	-0.0038	30.0	Yes
397.....	186	51,794	18.2698	0.0185	29.0	Yes
398.....	187	51,789	19.8354	0.0022	26.4	Yes
399.....	188	51,817	21.6700	0.0376	28.3	Yes
400.....	189	51,820	23.6367	0.0915	24.9	Yes
401.....	190	51,788	25.8247	-0.0208	23.2	No
402.....	191	51,793	27.9064	0.0018	28.6	Yes
403.....	192	51,871	29.8199	0.0102	24.8	Yes
404.....	193	51,812	31.6460	-0.0352	24.8	Yes
405.....	194	51,816	33.6829	-0.0329	20.2	Yes
406.....	195	51,817	35.8830	0.1250	30.4	Yes
407.....	196	51,820	37.8460	-0.0594	30.4	Yes
408.....	197	51,821	39.8224	0.0369	26.4	Yes
409.....	198	51,871	41.9984	0.0040	33.5	Yes
410.....	199	51,816	43.8616	-0.0170	24.0	Yes
411.....	200	51,817	45.7871	-0.0168	26.3	Yes
412.....	201	51,931	47.7341	0.0035	35.7	Yes
413.....	202	51,929	49.6868	0.0153	33.2	Yes
414.....	203	51,869	51.6314	-0.0118	39.7	Yes
415.....	204	51,810	53.5228	0.0307	25.7	No
416.....	205	51,811	55.4916	0.0138	38.6	Yes

NOTE.—Unobserved plates have no entry under the “Date Completed” column.

removed, “atlas” images, which include all significant pixels around each detected object, and a GIF image of each spectrum, with features identified;

*Spectra:* The flux- and wavelength-calibrated, sky-subtracted spectra, with error and mask arrays; and

*Other data products:* The number of objects loaded in the databases, summaries of observing conditions for imaging fields and for spectroscopic plates, and a variety of others.

### 2.2.1. Parameters, Classes, and Associations

We organize our measured parameters, for imaging and spectroscopic data, by grouping related parameters into *classes*. Table 5 lists the classes. Tables 6, 10, 12, 13, 14, 15, 16, 17, and 18 list the parameters in each class, with a brief description.

In each table, we list the parameter name (in many cases, parameters have several synonyms, as listed), the data type, a brief description, and, for Table 6, an indication of

whether this is a tagged entry (a parameter for which searches are particularly fast; see § 2.4.1). For binary flags, the meaning of the bits are given explicitly. Parameters for which there exists a placeholder in the database, but which are not yet calculated, are indicated as such with the phrase “placeholder” in the table.

Many quantities are defined for each band and thus have five entries. In the Catalog Archive Server (§ 2.3.1) these are indexed with 0, 1, 2, 3, and 4 inside of square brackets for *u*, *g*, *r*, *i*, and *z*, respectively. In the SkyServer (§ 2.3.3), these are indexed with *\_u*, *\_g*, *\_r*, *\_i*, and *\_z* for the filters.

Some parameters are *associations* to objects in other classes. For example, an object in the class SpecObj (Table 12) has a parameter *plate*, which is an object in the Plate class (Table 17). This contains information common to the 640 spectra taken in one set of observations for the plate. The association is indicated in the table in the type column as OneAssoc(Plate). Other associations point to many (rather than one) objects. In SpecObj, for example, the

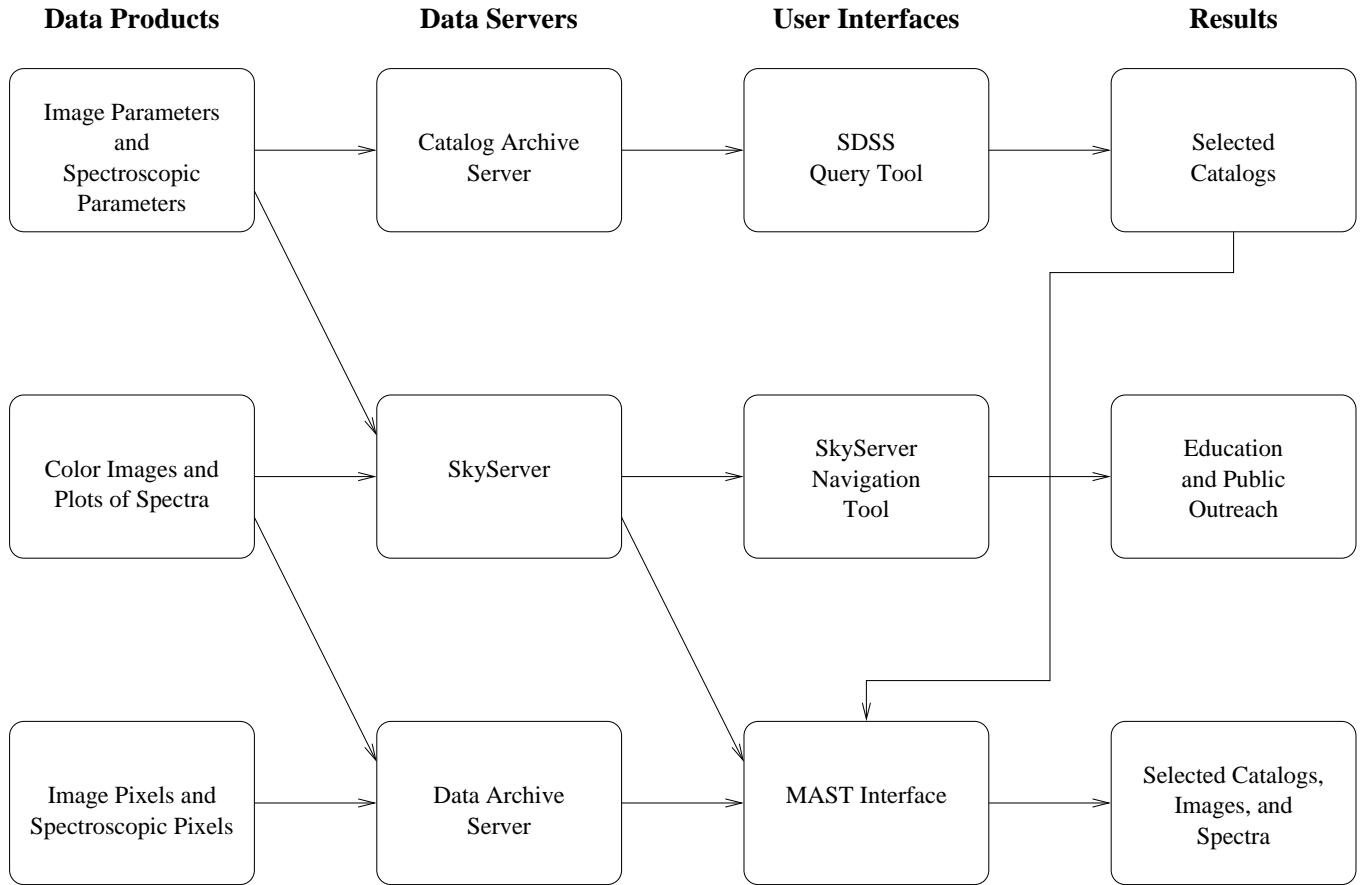


FIG. 3.—Overview of data products and distribution for the EDR. The left column contains all of the data products available. The second column contains the servers that hold data. Note that not all data products are contained in each of the servers. The third column contains the interfaces we provide to these servers. Choose which interface to use based on the results you wish to obtain, listed in the fourth column.

parameter `emissionredshift` has the type `ManyAssoc` (`EmissionRedshift`). It points to a list of redshifts measured. The “best” value of the emission and absorption redshift is stored in `z` of `SpecObj`, but for complex spectra, one may access the full list of redshifts measured during processing.

### 2.2.2. Imaging Data Products

#### 2.2.2.1. Image Parameters

The results of the imaging pipelines (described § 4.4) are summarized in Table 6, and the apertures used for radial profile measurements are in Table 7. For each object, we measure the position, several varieties of the flux, morphological parameters, and a provisional classification. We also include informational flags on each pixel in the images (Table 8) and on each object (Table 9) and summary flags and statistics for each field (Table 16). The choice of flags and which measure of flux and shape to use are dictated by the science goals of the SDSS. The catalog contains over 120 parameters and flags measured for 13,804,448 objects. Of these, 10,947,783 are unique, well-measured astrophysical objects.

The SDSS object catalogs are matched against the FIRST (Becker, White, & Helfand 1995), *ROSAT* (Voges et al. 1999), and USNO-A2.0 (Monet 1998) databases; the resulting parameters are described in Table 10.

#### 2.2.2.2. Images

Each corrected frame is a FITS image for one filter, 2048 columns by 1489 rows, with row number increasing in the scan direction. These are the imaging frames with flat-field, bias, cosmic-ray, and pixel-defect corrections applied. A raw image contains 1361 rows, and a corrected frame has the first 128 rows of the following corrected frame appended to it. The pixels subtend 0''.396 square on the sky. Header information using the world coordinate system (WCS) allows standard astronomical FITS tools to convert pixel position to right ascension and declination (§ 4.2.2).

#### 2.2.2.3. Binned Images

Each file is a FITS image for one filter,  $512 \times 372$  pixels, with WCS information. These are the corrected frames with detected objects removed and binned 4 pixels by 4 pixels.

#### 2.2.2.4. Mask Frame

Each file is a binary FITS table for one filter. Each row of the table describes a set of pixels in the corrected frame, using mask values described in Table 8.

#### 2.2.2.5. Atlas Image

For each detected object, the atlas image comprises the pixels that were detected as part of the object in any filter.

TABLE 5  
SUMMARY OF DATABASE CLASSES

Class	Description	Table
PhotoObj .....	Imaging parameters	6
PhotoPrimary .....	Primary detections	
PhotoSecondary .....	Secondary detections	
PhotoFamily .....	Bright or blended detections	
Tag .....	Tag imaging parameters	6
Primary .....	Primary detections	
Star .....	Primary star detections	
Galaxy .....	Primary galaxy detections	
Sky .....	Primary sky objects	
Unknown .....	Primary, unknown identification	
Secondary .....	Secondary detections	
StarSecondary .....	Secondary star detections	
GalaxySecondary .....	Secondary galaxy detections	
SkySecondary .....	Secondary sky objects	
UnknownSecondary .....	Secondary, unknown identification	
Family .....	Bright or blended detections	
ExternalCatalog .....	Objects matched from other catalogs	10
ROSAT .....	<i>ROSAT</i> sources	
FIRST .....	FIRST radio sources	
USNO .....	USNO-A2.0 sources	
SpecObj .....	Spectroscopic parameters	12
SpecLine .....	Spectroscopic lines	13
SpecLineIndex .....	Index for spectroscopic lines	13
CrossCorrelationRedshift .....	Measured cross-correlation redshifts	13
EmissionRedshift .....	Measured emission-line redshifts	13
Statistics .....	Number of entries in the database	14
Run .....	Statistics from one imaging run	15
Chunk .....	Statistics from interleaved pair of runs	15
Segment .....	One of 12 camera columns within a chunk	15
Field .....	Summary of measured parameters for one imaging field	16
Plate .....	Summary of parameters for one spectroscopic plate	17
SDSSConstants .....	Definition of the survey geometry	18

NOTE.—When indented, a class inherits all data members from the class listed above it.

These are provided through a database, as either a JPEG color image or a FITS file for each selected filter.

#### 2.2.2.6. Color Image

We combine the corrected frames from the  $g$ ,  $r$ , and  $i$  filters to produce a color image with the filters corresponding to blue, green, and red, respectively. The intensity mapping of each color is adjusted to enhance the appearance of these images. We use the same mapping for all of the color images.

#### 2.2.3. Reading SDSS Binary Tables

Most of these files are simple FITS binary tables or images; the exceptions are atlas images (*fpAtlas*), mask frames (*fpM*), and point-spread function description files (*fpField*). We provide stand-alone code on our Web sites to enable readers to interpret these files and read them into their own code.

#### 2.2.4. Spectroscopic Data Products

##### 2.2.4.1. Calibrated Spectra

These are the wavelength-flux pairs that constitute the spectrum. The spectra are given in vacuum wavelengths in the heliocentric frame, with flux density given in units of  $10^{-17}$  ergs  $s^{-1} cm^{-2} \text{\AA}^{-1}$ . An estimated error is also output,

as well as a mask associated with each pixel, as described in Table 11.

##### 2.2.4.2. Spectroscopic Parameters

Catalogs produced by the spectroscopic pipelines (described § 4.10) are summarized in Tables 12 and 13. For each spectrum, we measure the redshift using several techniques, locate and characterize lines, and assign an identification. The catalog contains 54,008 spectra, with 46 parameters measured for each spectrum, and 34 parameters measured for each emission line identified in each spectrum.

##### 2.2.4.3. Images of the Spectra

For convenience, we also provide a plot of each spectrum. These are GIF images of the spectrum, with significant features, our classification, and measured redshift indicated.

#### 2.2.5. Other Data Products

Table 14 summarizes parameters that record the number of objects loaded in the database. Table 15 summarizes data that describe the imaging data. A *run*, as already described, is all of the data from a single contiguous scan of the imaging camera. These are combined into *chunks*, contiguous areas of the sky in which spectroscopic targets will be selected and spectroscopic tiling (the process by which targets are assigned to spectroscopic plates) will be done. A

TABLE 6  
PARAMETERS ASSOCIATED WITH OBJECTS IN IMAGING DATA

Name	Type	Description	In Tag?
Class PhotoObj			
objID; id_; object_ID; obj_ID .....	int64	The unique identifier for SDSS photometric objects: a bit-encoded integer of run, rerun, camCol, field, objid	Yes
object; id; objid .....	int32	The object ID within a field (1-indexed). Usually changes between reruns of the same field	No
nChild; nchild; nChildren .....	int32	Number of children if a deblended object. BRIGHT objects also have nchild = 1, the non-BRIGHT sibling	Yes
objType; objc_type; objtype .....	int32	Classification of the object: 0 = OBJ_TYPE_UNK 1 = OBJ_TYPE_CR 2 = OBJ_TYPE_DEFECT 3 = OBJ_TYPE_GALAXY 4 = OBJ_TYPE_GHOST 5 = OBJ_TYPE_KNOWNOBJ 6 = OBJ_TYPE_STAR 7 = OBJ_TYPE_TRAIL 8 = OBJ_TYPE_SKY	Yes
catID; catid; cat_ID .....	int32	Catalog ID; not used	No
objFlags; objc_flags; flag1; objFlags2; objc_flags2; flag2 .....	int64	Detection flags; see Table 9	Yes
rowC; objc_rowc; objRowc .....	float32	Row center position, $r$ coordinates (pixels)	Yes
rowCErr; objc_rowcErr; objRowcErr .....	float32	Row center position error, $r$ coordinates (pixels)	No
colC; objc_colc; objColc .....	float32	Column center position, $r$ coordinates (pixels)	Yes
colCErr; objc_colcErr; objColcErr .....	float32	Column center position error, $r$ coordinates (pixels)	No
rowV; rowv .....	float32	Row component of object's velocity (deg day $^{-1}$ )	No
rowVErr; rowvErr .....	float32	Error in row component of object's velocity (deg day $^{-1}$ )	No
colV; colv .....	float32	Column component of object's velocity (deg day $^{-1}$ )	No
colVErr; colvErr .....	float32	Error in column component of object's velocity (deg day $^{-1}$ )	No
row; rowc .....	float32[5]	Row center (pixels)	No
rowErr; rowcErr .....	float32[5]	Row center error (pixels)	No
col; colc .....	float32[5]	Column center (pixels)	No
colErr; colcErr .....	float32[5]	Column center error (pixels)	No
sky .....	float32[5]	Sky flux at center of object (asinh mag arcsec $^{-2}$ )	No
skyErr .....	float32[5]	Sky flux error (asinh mag arcsec $^{-2}$ )	No
psfMag; psfCounts .....	float32[5]	PSF flux (asinh mag)	Yes
psfMagErr; psfCountsErr .....	float32[5]	PSF flux error (asinh mag)	No
fiberMag; fiberCounts .....	float32[5]	Flux in 3"-diameter fiber radius (asinh mag)	Yes
fiberMagErr; fiberCountsErr .....	float32[5]	Error in 3" fiber flux (asinh mag)	No
petroMag; petroCounts .....	float32[5]	Petrosian flux (asinh mag)	Yes
petroMagErr; petroCountsErr .....	float32[5]	Petrosian flux error (asinh mag)	No
petroRad .....	float32[5]	Petrosian radius (arcsec)	Yes
petroRadErr .....	float32[5]	Petrosian radius error (arcsec)	No
petroR50 .....	float32[5]	Radius containing 50% of Petrosian flux (arcsec). Only petroR50[2] is in Tag classes (Placeholder)	Yes
petroR50Err .....	float32[5]	Radius containing 90% of Petrosian flux (arcsec). Only petroR90[2] is in Tag classes (Placeholder)	No
petroR90 .....	float32[5]	Radius containing 90% of Petrosian flux (arcsec). Only petroR90[2] is in Tag classes (Placeholder)	Yes
petroR90Err .....	float32[5]	Stokes $Q$ -parameter	No
q; Q .....	float32[5]	Stokes $Q$ -parameter error	No
qErr; QErr; Q_err .....	float32[5]	Stokes $U$ -parameter	No
u; U .....	float32[5]		No

TABLE 6—Continued

Name	Type	Description	In Tag?
uErr; UErr; U_err .....	float32[5]	Stokes $U$ -parameter error	No
isoRowC; iso_rowc; isoRowc .....	float32[5]	Isophotal row centroid (pixels)	No
isoRowCErr; iso_rowcErr; isoRowcErr .....	float32[5]	Isophotal row centroid error (pixels) (placeholder)	No
isoRowCGrad; iso_rowcGrad; isoRowcGrad .....	float32[5]	Gradient in row centroid with isophote (pixels per 1 mag arcsec $^{-2}$ surface brightness change at object center)	No
isoColC; iso_colc; isoColc .....	float32[5]	Isophotal column centroid (pixels)	No
isoColCErr; iso_colcErr; isoColcErr .....	float32[5]	Isophotal column centroid error (pixels) (placeholder)	No
isoColCGrad; iso_colcGrad; isoColcGrad .....	float32[5]	Gradient in column centroid with isophote (pixels per 1 mag arcsec $^{-2}$ surface brightness change at object center)	No
isoA; iso_a .....	float32[5]	Isophotal major axis (arcsec)	No
isoAErr; iso_aErr .....	float32[5]	Isophotal major axis error (arcsec) (placeholder)	No
isoB; iso_b .....	float32[5]	Isophotal minor axis (arcsec)	No
isoBErr; iso_bErr .....	float32[5]	Isophotal minor axis error (arcsec) (placeholder)	No
isoAGrad; iso_aGrad .....	float32[5]	Gradient in major axis with isophote (arcsec per 1 mag arcsec $^{-2}$ surface brightness change at isophote)	No
isoBGrad; iso_bGrad .....	float32[5]	Gradient in minor axis with isophote (arcsec per 1 mag arcsec $^{-2}$ surface brightness change at isophote)	No
isoPhi; iso_phi .....	float32[5]	Isophotal position angle (deg + S through E)	No
isoPhiErr; iso_phiErr .....	float32[5]	Isophotal position angle error (deg) (placeholder)	No
isoPhiGrad; iso_phiGrad .....	float32[5]	Gradient in position angle with isophote (deg per 1 mag arcsec $^{-2}$ surface brightness change at isophote)	No
deVRad; r_deV .....	float32[5]	De Vaucouleurs fit scale radius (arcsec)	No
deVRadErr; r_deVErr .....	float32[5]	De Vaucouleurs fit scale radius error (arcsec)	No
deVaB; ab_deV .....	float32[5]	De Vaucouleurs fit $a/b$	No
deVaBErr; ab_deVErr .....	float32[5]	De Vaucouleurs fit $a/b$ error	No
deVPhi; phi_deV .....	float32[5]	De Vaucouleurs fit position angle (deg + E through N)	No
deVPhiErr; phi_deVErr .....	float32[5]	De Vaucouleurs fit position angle error (deg)	No
deVMag; counts_deV; deVCounts .....	float32[5]	De Vaucouleurs magnitude fit (asinh mag)	No
deVMagErr; counts_deVErr; deVCountsErr .....	float32[5]	De Vaucouleurs magnitude fit error (asinh mag)	No
expRad; r_exp .....	float32[5]	Exponential fit scale radius (arcsec)	No
expRadErr; r_expErr .....	float32[5]	Exponential fit scale radius error (arcsec)	No
expAB; ab_exp .....	float32[5]	Exponential fit $a/b$	No
expABErr; ab_expErr .....	float32[5]	Exponential fit $a/b$ error	No
expPhi; exp_phi .....	float32[5]	Exponential fit position angle (deg + E through N)	No
expPhiErr; exp_phiErr .....	float32[5]	Exponential fit position angle error (deg)	No
expMag; counts_exp; expCounts .....	float32[5]	Exponential fit (asinh mag)	No
expMagErr; counts_expErr; expCountsErr .....	float32[5]	Exponential fit error (asinh mag)	No
modelMag; counts_model; modelCounts .....	float32[5]	Better of de Vaucouleurs/exponential magnitude fit (asinh mag)	No
modelMagErr; counts_modelErr; modelCountsErr .....	float32[5]	Better of de Vaucouleurs/exponential magnitude fit error (asinh mag)	No
texture .....	float32[5]	Texture parameter (placeholder)	No
lStar; star_L .....	float32[5]	Star logarithmic likelihood; lStar_r, star_L_r in Tag classes	Yes
lExp; exp_L .....	float32[5]	Exponential disk fit logarithmic likelihood; lExp_r, exp_L_r in Tag classes	Yes
lDeV; deV_L .....	float32[5]	De Vaucouleurs fit logarithmic likelihood; lDeV_r, deV_L_r in Tag classes	Yes
fracPSF .....	float32[5]	Fraction of light in PSF (placeholder)	No
flags; flags2 .....	int64[5]	Same as objFlags, but for each band	No
type .....	int32[5]	Same as objType, but for each band	No

TABLE 6—Continued

Name	Type	Description	In Tag?
status .....	int32	Bit mask specifying the status of the object: AR_OBJECT_STATUS_SET Status has been set in reference to its own run AR_OBJECT_STATUS_GOOD Good as determined by its object flags; absence implies bad AR_OBJECT_STATUS_DUPLICATE Has one or more duplicate detections in an adjacent field of the same frames run AR_OBJECT_STATUS_OK_RUN Located in the primary range of rows for this field AR_OBJECT_STATUS_RESOLVED Has been resolved against other runs AR_OBJECT_STATUS_PSEGMENT Belongs to a primary segment; not necessarily a primary object AR_OBJECT_STATUS_FIRST_FIELD Belongs to the first field in its segment AR_OBJECT_STATUS_OK_SCANLINE Lies within valid $\nu$ range for its scan line AR_OBJECT_STATUS_OK_STRIPE Lies within valid $\eta$ range for its stripe AR_OBJECT_STATUS_SECONDARY This is a secondary survey object AR_OBJECT_STATUS_PRIMARY This is a primary survey object AR_OBJECT_STATUS_TARGET This is a spectroscopic target	Yes
ra.....	float64	J2000 right ascension, $\alpha$ ( $r$ ) (deg)	Yes
dec.....	float64	J2000 declination, $\delta$ ( $r$ ) (deg)	Yes
offsetRa.....	float32[5]	Filter position $\alpha$ minus final $\alpha \cos \delta$ (arcsec)	No
offsetDec.....	float32[5]	Filter position $\delta$ minus final $\delta$ (arcsec)	No
primTarget .....	int32	Bit mask specifying which primary target categories the object was selected in; see Table 27 for bit definitions	Yes
secTarget .....	int32	Bit mask specifying which secondary target categories the object was selected in; see Table 27 for bit definitions	Yes
reddening .....	float32[5]	Reddening in each filter (mag)	Yes
priority .....	uint32	Priority bits	No
rho.....	float32	Surface brightness parameter ( $5 \log \text{petroRad\_i}$ )	Yes
Database-specific Outputs			
cx, cy, cz .....	float64	Unit ( $x, y, z$ )-vector for $\alpha + \delta$	Yes
htmID.....	int64	Level 20 HTM <sup>a</sup> ID	Yes
modelFluxID .....	int32	$k-d$ tree ID for flux space (model flux)	Yes
psfFluxID .....	int32	$k-d$ tree ID for flux space (PSF flux)	Yes
petroFluxID .....	int32	$k-d$ tree ID for flux space (Petrosian flux)	Yes
matchID.....	int64[]	List of the IDs of the matching objects	No
field; Field.....	OneAssoc(Field)	Link to the field that this object is in	Yes
phototag; PhotoTag; tag .....	OneAssoc(PhotoTag)	Link to the Tag object corresponding to this object	No
profile; Profile.....	OneAssoc(Profile)	Link to the profile parameters for this object	Yes
externalcatalog; extern; ExternalCatalog .....	ManyAssoc(ExternalCatalog)	Link to external catalog associations, if any	Yes

TABLE 6—Continued

Name	Type	Description	In Tag?
specobj; SpecObj; spec	ManyAssoc(SpecObj)	Link to spectroscopic object associations, if any	Yes
child <sup>b</sup>	ManyAssoc(PhotoObj)	Link to the deblended children of this object, if any	Yes
parent <sup>b</sup>	OneAssoc(PhotoObj)	Link to the parent object if this object has been deblended	Yes
secondary <sup>b</sup>	ManyAssoc(PhotoObj)	Link to all secondaries of this object if it is a primary	Yes
primary <sup>b</sup>	ManyAssoc(PhotoObj)	Link to the primary object(s) if this object is a secondary	Yes
seclist <sup>b</sup>	ManyAssoc(PhotoObj)	Link to other secondary objects if this is a secondary	Yes
photoz; PhotoZ	OneAssoc(PhotoZ)	Link to the photometric redshift object (placeholder)	Yes
u <sup>c</sup>	float64	Better of de Vaucouleurs/exponential magnitude fit (asinh mag)	Yes
g <sup>c</sup>	float64	Better of de Vaucouleurs/exponential magnitude fit (asinh mag)	Yes
r <sup>c</sup>	float64	Better of de Vaucouleurs/exponential magnitude fit (asinh mag)	Yes
i <sup>c</sup>	float64	Better of de Vaucouleurs/exponential magnitude fit (asinh mag)	Yes
z <sup>c</sup>	float64	Better of de Vaucouleurs/exponential magnitude fit (asinh mag)	Yes

Class Profile: The Array of Profile Bins Calculated for Each Photometric Object

nProf; nprof	int32[5]	Number of radial bins	...
u; profMean_u	float32[]	Mean flux in annulus, <i>u</i> band (maggies arcsec <sup>-2</sup> )	...
g; profMean_g	float32[]	Mean flux in annulus, <i>g</i> band (maggies arcsec <sup>-2</sup> )	...
r; profMean_r	float32[]	Mean flux in annulus, <i>r</i> band (maggies arcsec <sup>-2</sup> )	...
i; profMean_i	float32[]	Mean flux in annulus, <i>i</i> band (maggies arcsec <sup>-2</sup> )	...
z; profMean_z	float32[]	Mean flux in annulus, <i>z</i> band (maggies arcsec <sup>-2</sup> )	...
uErr; profErr_u	float32[]	Error of mean flux in annulus, <i>u</i> band (maggies arcsec <sup>-2</sup> )	...
gErr; profErr_g	float32[]	Error of mean flux in annulus, <i>g</i> band (maggies arcsec <sup>-2</sup> )	...
rErr; profErr_r	float32[]	Error of mean flux in annulus, <i>r</i> band (maggies arcsec <sup>-2</sup> )	...
iErr; profErr_i	float32[]	Error of mean flux in annulus, <i>i</i> band (maggies arcsec <sup>-2</sup> )	...
zErr; profErr_z	float32[]	Error of mean flux in annulus, <i>z</i> band (maggies arcsec <sup>-2</sup> )	...
photoobj; PhotoObj; obj; pobj	OneAssoc(PhotoObj)	Link to the photometric object	...

<sup>a</sup> Hierarchical triangular mesh.<sup>b</sup> Data member is Assoc(PhotoTag) in Tag classes.<sup>c</sup> Magnitude appears only in PhotoTag classes.

TABLE 7  
RADII FOR profMean AND profErr  
MEASUREMENTS

Aperture	Radius (pixels)	Radius (arcsec)	Area (pixels)
1.....	0.56	0.23	1
2.....	1.69	0.68	9
3.....	2.58	1.03	21
4.....	4.41	1.76	61
5.....	7.51	3.00	177
6.....	11.58	4.63	421
7.....	18.58	7.43	1085
8.....	28.55	11.42	2561
9.....	45.50	18.20	6505
10.....	70.15	28.20	15619
11.....	110.50	44.21	38381
12.....	172.50	69.00	93475
13.....	269.50	107.81	228207
14.....	420.50	168.20	555525
15.....	657.50	263.00	1358149

*segment* is a single piece of a run for a single camera column. The contiguous stream of data is divided into a series of *fields* (§ 3.5), whose detailed properties are also given in Table 16, including astrometric calibrations (§ 4.2) and details of the point-spread function fitting (§ 4.3). Finally, the details for each spectroscopic plate are also given in Table 17.

Table 18 lists constants we use to define the survey.

### 2.3. Database Servers

We have constructed three database servers to hold the imaging and spectroscopic data for the EDR, as shown in Figure 3. The *Catalog Archive Server* contains the measured parameters from all objects in the imaging survey and the spectroscopic survey. These are loaded in a database server that we built using Objectivity/DB, a commercial, object-oriented database server (Objectivity 2001). The *SkyServer* contains identical information, but it is loaded in a database server we built using a commercial, relational database server, SQL Server (Delaney 2000). The *Data Archive Server* contains the rest of the data products for the EDR, such as the corrected imaging frames, which are available for direct download, and the calibrated spectra, which are loaded in an Objectivity database.

### 2.3.1. Catalog Archive Server

The Catalog Archive Server is indexed to allow efficient queries on quantities commonly used in astronomical research (position, and magnitudes in our five filters). It also allows algebraic combinations of these quantities to be used in the queries. We have written a unique and powerful database server that allows users to make sophisticated queries efficiently. The general strategy of compiling a large astronomical database is described in Szalay et al. (2001), and our method of dividing the celestial sphere into a hierarchical triangular mesh, which enables efficient access based on position, is described in Kunszt et al. (2000). The database server software is described in Thakar et al. (2000).

### 2.3.2. Data Archive Server

Given a set of object coordinates, either from the Catalog Archive Server or some other source, the Data Archive Server makes the detailed data (corrected frames, binned images, mask images, atlas images, color images, spectra, and spectral plots) available. It is not practical to access all of the imaging data in this way, but it does give convenient access to any selected field or spectrum contained in the EDR.

### 2.3.3. SkyServer

The SkyServer is a relational database server. Its language is not as rich as that provided by the Catalog Archive Server, but it works well for many queries. This server was originally developed with outreach and education in mind, but it is also very useful for astronomical research.

## 2.4. User Interfaces

There are three separate user interfaces to the SDSS EDR, as shown in Figure 3. We expect that typical users interested in extracting subsets of our imaging catalog based on position, flux, and colors will find the interface to the SkyServer at the Multimission Archive at STScI (MAST) most useful. Sophisticated queries on flags and algebraic combinations of parameters are best done with the SDSS Query Tool (sdssQT). Users who wish to download atlas images, corrected frames, or spectra should do so with the MAST interface to the Data Archive Server.

### 2.4.1. SDSS Query Tool (sdssQT)

The SDSS Query Tool is a stand-alone application to manage and perform Catalog Archive Server queries and is

TABLE 8  
IMAGE MASK BITS

Hexadecimal Bit	Flag	Explanation
0x001 .....	INTERP	Pixel's value has been interpolated
0x002 .....	SATUR	Pixel is/was saturated
0x004 .....	NOTCHECKED	Pixel was <i>not</i> examined for an object
0x008 .....	OBJECT	Pixel is part of some object
0x010 .....	BRIGHTOBJECT	Pixel is part of bright object
0x020 .....	BINOBJECT	Pixel is part of binned object
0x040 .....	CATOBJECT	Pixel is part of a catalogued object
0x080 .....	SUBTRACTED	Model has been subtracted from pixel
0x100 .....	HOST	Pixel is part of a ghost
0x200 .....	CR	Pixel is part of a cosmic ray

TABLE 9  
PROCESSING FLAGS ASSOCIATED WITH MEASURING OBJECTS IN frames

Name	Hexadecimal	Bit	Description
BRIGHT .....	0x2		Object was detected in first, “bright” object finding step; generally brighter than $r^* = 17.5$
SUBTRACTED .....	0x100000		This BRIGHT object had its wings subtracted from the frame
BINNED1 .....	0x10000000		Object was detected in the $1 \times 1$ , unbinned image
BINNED2 .....	0x20000000		Object was detected in the $2 \times 2$ binned image, after unbinned detections are replaced by background
BINNED4 .....	0x40000000		Object was detected in the $4 \times 4$ binned image
EDGE .....	0x4		Object was too close to edge of frame
LOCAL_EDGE .....	0x8000000000		Center in at least one band is too close to an edge
SATUR .....	0x40000		Object contains one or more saturated pixels
SATUR_CENTER .....	0x800000000000		Object center is close to at least one saturated pixel
BLENDED .....	0x8		Object had multiple peaks detected within it and was thus a candidate to be a deblending parent
NODEBLEND .....	0x40		No deblending was attempted on this object, even though it is BLENDED
DEBLEND_TOO_MANY_PEAKS .....	0x800		There were more than 25 peaks in this object to deblend; deblended brightest 25
DEBLEND_PRUNED .....	0x4000000		At least one child was removed because its image was too similar to a supposedly different child
PEAKS_TOO_CLOSE .....	0x2000000000		At least some peaks within this object were too close to be deblended, and thus they were merged into a single peak
DEBLEND_UNASSIGNED_FLUX .....	0x4000000000		After deblending, a significant fraction of flux was not assigned to any children
DEBLENDED_AT_EDGE .....	0x200000000000		An object close enough to the edge of the frame that it normally would not be deblended is deblended anyway. Only set for objects large enough to be EDGE in all fields/strips
CHILD .....	0x10		Object is the product of an attempt to deblend a BLENDED object
DEBLENDED_AS_PSF .....	0x2000000		The deblender treated this object as a PSF
DEBLEND_NOPEAK .....	0x400000000000		There was no detected peak within this child in at least one band
MOVED .....	0x80000000		The deblender identified this object as possibly moving
DEBLENDED_AS_MOVING .....	0x100000000000		A MOVED object that the deblender treated as moving
NODEBLEND_MOVING .....	0x200000000000		A MOVED object that the deblender did not treat as moving
TOO_FEW_DETECTIONS .....	0x4000000000		A child of this object was not detected in enough bands to reliably deblend as moving
BAD_MOVING_FIT .....	0x8000000000		The moving fit was too poor to be believable
STATIONARY .....	0x100000000000		Object was consistent with being stationary
BAD_MOVING_FIT_CHILD .....	0x200000000000		A possible moving child’s velocity fit was too poor, so it was discarded and the parent was not deblended as moving
TOO_FEW_GOOD_DETECTIONS .....	0x10000000000000		A child of this object had too few good detections to be deblended as moving
PEAKCENTER .....	0x20		Given center is position of peak pixel, rather than based on the maximum likelihood estimator
CANONICAL_CENTER .....	0x1		Measurements used the center in $r$ , rather than the locally determined center
NOTCHECKED .....	0x80000		There are pixels in the object that were not checked to see whether they included a local peak, such as cores of saturated stars
MEDIAN_CENTRE .....	0x400000000000		Center given is of median-smoothed image
CENTER_OFF_AIMAGE .....	0x20000000000000		At least one peak’s center lay off of the atlas image. This can happen when the object is deblended as moving, or if the astrometry is bad
NOPETRO .....	0x100		No valid Petrosian radius was found for this object
MANYPETRO .....	0x200		More than one Petrosian radius was found
NOPETRO_BIG .....	0x400		Petrosian radius is beyond the last point in the radial profile
PETROFAINT .....	0x800000		At least one possible Petrosian radius was rejected, as the surface brightness at $r_p$ was too low. If NOPETRO is not set, a different, acceptable Petrosian radius was found
MANYR50 .....	0x2000		Object has more than one 50% light radius
MANYR90 .....	0x4000		Object has more than one 90% light radius
BADSKY .....	0x400000		The sky level is so bad that the highest pixel in the object is <i>very</i> negative, far more so than a mere nondetection. No further analysis is attempted
NOPROFILE .....	0x80		Object was too small or too close to the edge to estimate a radial profile
TOO_LARGE .....	0x1000000		Object is too large to measure its profile (i.e., it extends beyond a radius of approximately $260''$ ), or at least one child is larger than half a frame
BAD_RADIAL .....	0x8000		Some of the points in the given radial profile have negative S/N. Not a significant parameter
INCOMPLETE_PROFILE .....	0x10000		Petrosian radius intersects the edge of the frame
ELLIPFAINT .....	0x8000000		Object’s center is fainter than the isophote whose shape is desired, so the isophote properties are not measured. Also flagged if profile is incomplete
CR .....	0x1000		Contains a pixel interpreted to be part of a cosmic ray
INTERP .....	0x20000		The object contains one or more pixels whose values were determined by interpolation
INTERP_CENTER .....	0x100000000000		Object center is close to at least one interpolated pixel
PSF_FLUX_INTERP .....	0x800000000000		Greater than 20% of the PSF flux is from interpolated pixels
BAD_COUNTS_ERROR .....	0x100000000000		An object containing interpolated pixels had too few good pixels to form a reliable estimate of its error; the quoted error may be underestimated
NOSTOKES .....	0x200000		No Stokes parameter was measured, because of either an incomplete profile or numerical difficulties

NOTE.—These flags are set both in each band and, in some cases, for the object as a whole. A more thorough description of the flags and their meanings is available on our Web sites.

TABLE 10  
IMAGE PARAMETERS FOR OBJECTS FROM EXTERNAL CATALOGS

Name	Type	Description
Class ExternalCatalog: External Object		
delta.....	float32	Position difference between catalog and <i>r</i> Photo object (arcsec)
match.....	int32	Match to different catalogs: USNO, 1 = match to catalog within 30" FIRST, number of matches to catalog within 1"5 ROSAT, 1 = match to catalog within 60"
type.....	int16	Object match type: ROSAT_MATCH = 0 FIRST_MATCH = 1 USNO_MATCH = 2
phototag; PhotoTag; obj; pobj .....	OneAssoc(PhotoTag)	Link to PhotoTag object
Class FIRST: SDSS Objects That Match to FIRST Objects <sup>a</sup>		
peak; firstPeak.....	float32	Peak FIRST radio flux (mJy)
integr; int; firstInt .....	float32	Integrated FIRST radio flux (mJy)
rms; firstRms.....	float32	rms error in flux (mJy)
major; firstMajor .....	float32	Major axis (deconvolved) (arcsec)
minor; firstMinor .....	float32	Minor axis (deconvolved) (arcsec)
pa; firstPa .....	float32	Position angle (east of north) (deg)
id; firstId .....	int32	FIRST catalog ID of one matching source
lambda.....	float64	Composite radio source survey longitude (deg)
eta.....	float64	Composite radio source survey latitude (deg)
Class ROSAT: SDSS Objects That Match to ROSAT Objects <sup>a</sup>		
cps; rosatCps.....	float32	Integrated ROSAT counts (counts s <sup>-1</sup> )
hard1; rosatHr1; hr1 .....	float32	ROSAT hardness ratio 1
hard2; rosatHr2; hr2 .....	float32	ROSAT hardness ratio 2
extent; rosatExt; ext .....	float32	ROSAT extent parameter (arcsec)
exposure; rosatExposure .....	float32	ROSAT exposure time (seconds)
posErr; rosatPosErr; poserr.....	float32	Error in ROSAT position (arcsec)
cpsErr; rosatCpsErr; cpserr.....	float32	Error in integrated ROSAT counts (counts s <sup>-1</sup> )
hard1Err; rosatHr1Err; hr1Err ....	float32	Error in ROSAT hardness ratio 1
hard2Err; rosatHr2Err; hr2Err ....	float32	Error in ROSAT hardness ratio 2
extentLike; rosatExtLike .....	float32	ROSAT extent parameter likelihood
detectLike; rosatDetectLike.....	float32	ROSAT detection likelihood
Class USNO: SDSS Objects That Match to USNO-A2.0 Objects <sup>a</sup>		
propermotion .....	float32	Proper motion of object (arcsec per century)
angle; propermotionangle .....	float32	Direction of proper motion (deg + N through E)
blue; usnoBlue.....	float32	Blue magnitude of matching USNO object (mag)
red; usnoRed.....	float32	Red magnitude of matching USNO object (mag)

<sup>a</sup> The SDSS objects have their match parameters stored here; each class inherits everything from ExternalCatalog, plus the parameters listed beneath.

available for download from our Web sites, along with a detailed user's manual. The discussion that follows is a brief introduction to its use, to illustrate its capabilities. There are options to choose the output format (ASCII or binary), facilities to save queries, a simple text editor to create and modify queries, and communication with the Catalog Archive Server to follow the progress (and predict the time to completion) of active queries. An additional tool converts our binary output format to a FITS binary table.

The query language we developed is very similar to the Structured Query Language (SQL). The Query Tool includes several example queries, and the on-line user's guide provides additional explanations of the language and how to use it efficiently. The use of associations (§ 2.2.1) provides a powerful way to extract object data from many different classes simultaneously. Similarly, the inheritance

properties of classes and their subclasses makes queries written for a given class run on all of its subclasses or sibling classes.

The grammar of this language is to select a set of *parameters* from a *class* that satisfy specified *conditions*. The SDSS Query Tool allows full access to all of the classes and parameters in the Catalog Archive Server. For example, the query

```
SELECT ra,dec FROM SpecObj WHERE (z > 2)
```

returns the parameters *ra* and *dec* (right ascension and declination) for all spectroscopic objects (i.e., from the class SpecObj) with redshift greater than 2.

One class must appear in the *FROM* clause. One or more of the parameters from the class may be listed in the *SELECT* clause and used in the *WHERE* clause, as long as they are included in the class mentioned in the *FROM* clause.

TABLE 11  
SPECTROSCOPIC MASK BITS

Hexadecimal Bit	Flag	Explanation
Flags Set on the Entire Spectrum		
0x001 .....	NOPLUG	Fiber not listed in plug-map file
0x002 .....	BADTRACE	Bad trace of flat field
0x004 .....	BADFLAT	Low counts in fiber flat
0x008 .....	BADARC	Bad arc solution
0x010 .....	MANYBADCOLUMNS	More than 10% of pixels are bad columns
0x020 .....	MANYREJECTED	More than 10% of pixels are rejected in extraction
0x040 .....	LARGESHIFT	Large spatial shift between flat and object position
0x080 .....	BADSKYFIBER	Sky fiber shows extreme residuals
0x100 .....	NEARWHOPPER	Within two fibers of a very bright object
0x400 .....	SMEARIMAGE	Smear available for red and blue cameras
0x800 .....	SMEARHIGHSN	S/N sufficient for full smear fit
0x1000 .....	SMEARMEDSN	S/N only sufficient for scaled median fit
Flags Set on Single Pixels		
0x10000.....	NEARBADPIXEL	Bad pixel within 3 pixels of trace
0x20000.....	LOWFLAT	Flat field less than 0.5
0x40000.....	FULLREJECT	Pixel fully rejected in extraction
0x80000.....	PARTIALREJECT	Some pixels rejected in extraction
0x100000.....	SCATTEREDLIGHT	Scattered light significant
0x200000.....	CROSSTALK	Cross talk significant
0x400000.....	NOSKY	Sky level unknown at this wavelength
0x800000.....	BRIGHTSKY	Sky level $>$ flux + $10 \times$ flux error and sky $>$ $1.25 \times$ median sky over 99 pixels
0x1000000.....	NODATA	No data available in combined B-spline
0x2000000.....	COMBINEREJ	Rejected in combined B-spline
0x4000000.....	BADFLUXFACTOR	Low flux calibration or flux correction factor
0x8000000.....	BADSKYCHI	$\chi^2 > 4$ in sky residuals at this wavelength
0x10000000.....	REDMONSTER	Contiguous region of bad $\chi^2$ in sky residuals

The PhotoObj class, containing the detected objects in the images, has the most entries, over 13 million. To facilitate selecting from these entries, we have designated a subset of the most commonly used parameters in the PhotoObj class to be part of a special class called Tag. The final column of Table 6 indicates which parameters are in the Tag class. The database is structured in such a way that searches that select on these Tag parameters run significantly faster.

In addition, many of the classes, such as Tag, have numerous subclasses (Primary, Secondary, Galaxy, Star), described in the aforementioned tables. These subclasses all *inherit* the properties of the umbrella class. The same parameters are available in the subclasses as in the umbrella class, but they will be faster to query, as each contains only a portion of the total objects. For example, users wanting data only on galaxies can execute queries on the Galaxy class; this avoids having to specify that objType = Galaxy if it were run on the entire Tag class and will run faster, as fewer objects are searched.

To access *associated*, or linked parameters, we use the “.” modifier. For example, spectroscopic objects from the SpecObj class all have a link back to the photometric object in the Tag class. This allows retrieval of parameters from both classes simultaneously. Use the syntax tag.r, for example, to obtain a spectroscopic target’s  $r^*$  magnitude. In Table 12, the parameter phototag has type OneAssoc(PhotoTag), so the query

```
SELECT ra,dec,phototag.r FROM SpecObj
```

returns the position and  $r^*$  magnitude for all objects with spectra.

Some associations tie many objects in a class to one object of another class. For example, there can be many lines measured in one spectrum. The parameter measured in the SpecObj class has the type ManyAssoc(SpecLine). The following *association query* returns parameters for all of the lines of the selected spectrum:

```
SELECT name.name, wave, ew FROM
  (SELECT measured FROM SpecObj WHERE
    plate.plateID == 384 && fiberID == 284)
```

The on-line user’s guide to the Query Tool contains a detailed description of the language, examples to help construct advanced queries, and details about macros to perform logical and arithmetic operations within queries.

#### 2.4.2. MAST SDSS SkyServer Interface

The STScI SkyServer interface provides a familiar way to access the object catalog, as it is based on the MAST interface,<sup>63</sup> which also serves catalogs from NASA missions and other surveys: the Guide Star Catalog, the Digitized Sky Survey, and the VLA FIRST survey. This is the interface most will prefer for straightforward queries based on right ascension and declination. Objects that satisfy the query can be used one of several ways. The list may be written to a comma-delimited text file, to be read by external programs, such as Microsoft Excel. The list may also be written as a FITS binary table. Finally, the list may be browsed as an HTML table, where specific objects and data products (e.g.,

<sup>63</sup> See <http://archive.stsci.edu/index.html>.

TABLE 12  
PARAMETERS DERIVED FROM SPECTRA

Name	Type	Description
Class SpecObj: Spectroscopic Object		
specobj_ID; spec_ID.....	int64	Unique ID
xFocal.....	float64	X focal position on plate (mm)
yFocal.....	float64	Y focal position on plate (mm)
fiberID.....	int16	Fiber ID (1–640)
z.....	float32	Final redshift (heliocentric)
zErr; z_err .....	float32	Redshift error
zConf; z_conf.....	float32	Redshift confidence (between 0 and 1)
zStatus; z_status.....	int16	Redshift status: 0 = NOT_MEASURED Not yet measured (sky or broken fiber) 1 = FAILED Redshift measurement failed; CL < 0.35 2 = INCONSISTENT Cross-correlation and emission-line redshifts both CL > 0.75 but differ by > 0.1 3 = XCORR_EMLINE Cross-correlation plus consistent emission-line redshift measurement 4 = XCORR_HIC z determined from cross-correlation with CL > 0.75 5 = XCORR_LOC z determined from cross-correlation with CL < 0.75 6 = EMLINE_XCORR Emission-line z plus consistent cross-correlation redshift measurement 7 = EMLINE_HIC z determined from emission lines with CL > 0.75 8 = EMLINE_LOC z determined from emission lines with CL < 0.75 9 = MANUAL_HIC z determined “by hand” with high confidence 10 = MANUAL_LOC z determined “by hand” with low confidence 11 = XCORR_4000BREAK Cross-correlation redshift determined when EW(4000 Å break) > 0.95
specClass; spec_cln.....	int16	Spectral classification: 0 = SPEC_UNKNOWN                          Unknown 1 = SPEC_STAR                              Star 2 = SPEC_GALAXY                           Galaxy 3 = SPEC_QSO                               Quasar 4 = SPEC_HIZ_QSO High-redshift quasar, z > 2.3; triggers use of the Ly $\alpha$ finding code 5 = SPEC_SKY                               Sky 6 = STAR_LATE Star dominated by molecular bands; M or later 7 = GAL_EM                                   Emission-line galaxy (placeholder)

TABLE 12—Continued

Name	Type	Description
zWarning; z_warning.....	uint32	Bit mask for warning flags: 0x00000 = Z_WARNING_OK No warnings 0x00001 = Z_WARNING_NO_SPEC No spectrum 0x00002 = Z_WARNING_NO_BLUE No blue-side spectrum 0x00004 = Z_WARNING_NO_RED No red-side spectrum 0x00010 = Z_WARNING_NOT_GAL Classification does not match galaxy target 0x00020 = Z_WARNING_NOT_QSO Classification does not match quasar target 0x00040 = Z_WARNING_NOT_STAR Classification does not match star target 0x00080 = Z_WARNING_GAL_COEF Galaxy PCA coefficients far off locus (placeholder) 0x00100 = Z_WARNING_EMAB_INC Emission and absorption redshifts inconsistent 0x00200 = Z_WARNING_AB_INC Absorption redshifts inconsistent 0x00400 = Z_WARNING_EM_INC Emission redshifts inconsistent 0x00800 = Z_WARNING_HIZ Redshift is high 0x01000 = Z_WARNING_LOC Confidence is low 0x02000 = Z_WARNING_LOW_SNG S/N is low in <i>g</i> 0x04000 = Z_WARNING_LOW_SNR S/N is low in <i>r</i> 0x08000 = Z_WARNING_LOW_SNI S/N is low in <i>i</i> 0x10000 = Z_WARNING_4000BREAK EW(4000 Å break) > 0.95
velDisp; vel_dis.....	float32	Velocity dispersion (km s <sup>-1</sup> ) (placeholder)
velDispErr; vel_dise.....	float32	Velocity dispersion error (km s <sup>-1</sup> ) (placeholder)
eCoeff; ecoeff.....	float32[5]	Expansion coefficients for galaxies
eClass; eclass.....	float32	Classification
nGood.....	int32	Number of good pixels in spectrum
sn.....	float32[3]	Median S/N in <i>g, r, i</i>
mag.....	float32[3]	Synthetic magnitude in <i>g*</i> , <i>r*</i> , <i>i*</i>
ra.....	float64	Right ascension (deg)
dec.....	float64	Declination (deg)
htmID.....	int64	HTM ID
fiberCounts .....	float32[5]	Fiber magnitudes (asinh mag)
objID.....	int64	SDSS ID of object
primTarget .....	int32	Primary-target flag (Table 28)
secTarget .....	int32	Secondary-target flag
pixMin.....	int16	Minimum pixel number
pixMax.....	int16	Maximum pixel number
objType .....	char[32]	Object type as string
phototag; PhotoTag; obj: tag .....	OneAssoc(PhotoTag)	Link to corresponding photometric object
emissionredshift; elz; emz; EmissionRedshift .....	ManyAssoc(EmissionRedshift)	Link to the emission-line redshift measures
crosscorrelationredshift; CrossCorrelationRedshift; xcorrz; xcz.....	ManyAssoc(CrossCorrelationRedshift)	Link to the redshifts from cross-correlations
found.....	ManyAssoc(SpecLine)	Link to found spectral lines
measured.....	ManyAssoc(SpecLine)	Link to measured spectral lines
speclineindex; lineidx; SpecLineIndex.....	ManyAssoc(SpecLineIndex)	Link to spectral line indices
plate; Plate.....	OneAssoc(Plate)	Link to corresponding plate object

TABLE 13  
SPECTROSCOPIC PARAMETERS: LINES AND REDSHIFTS

Name	Type	Description
Class SpecLine: Identified Spectroscopic Lines (In Observed Frame)		
specline_ID .....	int64	Unique ID
wave .....	float32	Line center (Å)
waveErr .....	float32	Error in line center (Å)
waveMin .....	float32	Minimum wavelength of line influence (Å)
waveMax .....	float32	Maximum wavelength of line influence (Å)
sigma .....	float32	Width ( $\sigma$ ) of fitted Gaussian (Å)
sigmaErr .....	float32	Error of $\sigma$ (Å)
sigmaMin .....	float32	Minimum $\sigma$ allowed by fit (Å)
sigmaMax .....	float32	Maximum $\sigma$ allowed by fit (Å)
height .....	float32	Height of Gaussian ( $10^{-17}$ ergs s $^{-1}$ cm $^{-2}$ Å $^{-1}$ )
heightErr .....	float32	Error in height ( $10^{-17}$ ergs s $^{-1}$ cm $^{-2}$ Å $^{-1}$ )
continuum .....	float32	Continuum value at this pixel ( $10^{-17}$ ergs s $^{-1}$ cm $^{-2}$ Å $^{-1}$ )
ew .....	float32	Equivalent width (Å)
ewErr .....	float32	Error in equivalent width (Å)
ewMin .....	float32	Minimum equivalent width allowed for line detection (Å)
specIndex .....	float32	Spectral index (not used)
nSigma .....	float32	Significance of line number of $\sigma$ (not used)
chisq .....	float32	$\chi^2$ of fit
nu .....	int16	Number of degrees of freedom
restWave .....	float32	Line rest wavelength (Å)
weight .....	float32	Weight of line in redshift determination
z .....	float32	Redshift of line (heliocentric)
zErr .....	float32	Error in redshift
lineMask .....	uint32	Line mask (not used)
category .....	int16	Category: line found (0) or measured (1)
specobj; spec .....	OneAssoc(SpecObj)	Link to spectrum object
speclinename; linename; name .....	OneAssoc(SpecLineName)	Link to line name
Class SpecLineName: Names of Spectral Lines		
lineID .....	int32	Line ID
restWave .....	float32	Line rest wavelength (Å)
name .....	char[16]	Line name
Class SpecLineIndex: Spectral Indices (In Rest Frame)		
speclineindex_ID .....	int64	Unique ID
name .....	char[16]	Name
ew .....	float32	Equivalent width (Å) or flux ratio
ewErr .....	float32	Equivalent width error (Å)
mag .....	float32	Magnitude (2.5 log EW) of the line index
magErr .....	float32	Magnitude error
waveMin .....	float32	Minimum rest wavelength used to calculate the index
waveMax .....	float32	Maximum rest wavelength used to calculate the index
z .....	float32	Redshift
sn .....	float32	Signal-to-noise ratio
specobj; spec .....	OneAssoc(SpecObj)	Link to spectrum object
Class CrossCorrelationRedshift: Results of Cross-Correlation Redshift Measurement		
crosscorrelationredshift_ID; xcz_ID .....	int64	Unique ID
tempNo; temp_no .....	int16	Template number
peakNo; peak_no .....	int16	Peak number
shift .....	float32	Number of pixels between object and template
z .....	float32	Redshift from this peak
zErr; z_err .....	float32	Error in redshift
r .....	float32	Tonry-Davis $r$
zConf; con; conf; z_conf .....	float32	Confidence
height .....	float32	Height of CCF peak
width .....	float32	Width of CCF peak
widthErr; width_err .....	float32	Error in CCF width
delta .....	float32	Error in CCF peak
rms .....	float32	rms of CCF
specobj; spec .....	OneAssoc(SpecObj)	Link to spectrum object
Class EmissionRedshift: Results of Emission-Line Redshift Measurement		
emissionredshift_ID; emz_ID .....	int64	Unique ID
z .....	float32	Redshift
zErr .....	float32	Error in redshift
zConf .....	float32	Confidence in redshift
nFit .....	int32	Number of matched lines
wtFit; fitWeight .....	float32	Weight of fit
specobj; spec .....	OneAssoc(SpecObj)	Link to spectrum object

TABLE 14  
STATISTICS CLASS

Name	Type	Description
Class Statistics: Total Number of Chunks, Runs, Fields, Objects, and Segments		
nChunk .....	int32	Number of chunk objects
nRun .....	int32	Number of run objects
nSegment .....	int32	Number of segments
nField .....	int32	Number of fields
nPhotoObj .....	int32	Number of Photo objects
nPrimary .....	int32	Number of primary Photo objects
nSecondary .....	int32	Number of secondary Photo objects
nFamily .....	int32	Number of Family objects
nStar .....	int32	Number of objects classified as stars
nGalaxy .....	int32	Number of objects classified as galaxies
nSky .....	int32	Number of objects classified as sky
nUnknown .....	int32	Number of objects classified as unknown
nStarSec .....	int32	Number of secondary objects classified as stars
nGalaxySec .....	int32	Number of secondary objects classified as galaxies
nSkySec .....	int32	Number of secondary objects classified as sky
nUnknownSec .....	int32	Number of secondary objects classified as unknown
nExternal .....	int32	Number of external objects
nPlate .....	int32	Number of plates
nSpecObj .....	int32	Number of spectra
nSpecLine .....	int32	Number of spectral lines
nSpecLineIndex.....	int32	Number of spectral line indices

the corrected frame for the object, or the spectra for those objects with spectra) are then selected for use by the Data Archive Server.

#### 2.4.3. MAST SDSS Data Archive Server Interface

The *SDSS Image Products* interface accepts a list of objects specified in one of the following formats: run/rerun/camCol/fieldID/objectID (see § 4.4.1), long object ID (from the SkyServer or Catalog Archive Server), equatorial coordinates, and NED or SIMBAD target name. Files generated by *sdssQT* may be uploaded directly. All imaging products (atlas images, corrected frames, reconstructed frames, binned images, fields summaries, and mask files) are available. All of the files for a request are bundled together in a .tar, .tar.gz, or .zip file.

An object list from the MAST SkyServer interface can be sent to the data product interfaces via a “shopping cart.” To demonstrate this access, we begin in the MAST SDSS SkyServer interface. Select galaxies within  $10'$  of  $(\alpha, \delta) = (0, 0)$  and  $16 < r^* < 17$  by entering the coordinates in the “RA” and “Dec” boxes and 10.0 in the “Radius” box, highlighting GALAXY from the object type list, and by entering 16..17 in the “Magnitudes” box for  $r$ . Three objects are returned. Choosing the “Browse Results as HTML” button brings up a list of these objects. Select the three objects by using the button in the “Mark” column, select “Add marked records to shopping cart,” and then select “Retrieve data products for shopping cart.” The data products may then be downloaded.

#### 2.5. Access to Files

Data products are also available directly from the Data Archive Server. Sophisticated users may want to download some of these files directly, but we consider these files to be intermediate data products and focus our support and documentation efforts on the database servers and interfaces we

described above. The data model, which describes the file naming conventions and FITS headers, is available on our Web sites.

#### 2.5.1. Imaging Files

Image data are available from the main Data Archive Server page. Data for individual runs are in subdirectories specified by \$run/\$rerun, where valid \$run/\$rerun combinations for the EDR are 94/7, 125/7, 752/8, 756/8, 1336/2, 1339/2, 1356/2, and 1359/3. Corrected frames are named as

```
corr/$camCol/
fpC-$run-$filter$camCol-$field.fit.gz,
```

where \$camCol is the camera column (1 through 6), \$run is the run number (six digits with leading zeroes), \$filter is the filter name (u, g, r, i, z), and \$field is the field number (four digits with leading zeros). Binned images are accessed in a similar fashion, via

```
objcs/$camCol/
fpBIN-$run-$filter$camCol-$field.fit.gz,
```

as are the mask images, via

```
objcs/$camCol/
fpM-$run-$filter$camCol-$field.fit.
```

The fpAtlas files in that same directory contain all of the atlas images for each field, but it is generally more convenient to access them with the Data Archive Server.

The catalogs of detected objects for each field are accessed via

```
objcs/$camCol/
fpObjc-$run-$camCol-$field.fit,
```

but these files do not yet include photometric or astrometric calibration. Summary information for each field processed

TABLE 15  
IMAGING RUN, CHUNK, AND SEGMENT CLASSES

Name	Type	Description
Class Run: One Continuous Photometric Observation		
run; runNumber.....	int32	Run number
skyNodes .....	int64[]	Level 5 HTM nodes of this run
chunk; Chunk.....	ManyAssoc(sxChunk)	Link to the chunks this run is part of
segment; Segment.....	ManyAssoc(sxSegment)	Link to the segments of this run
Class Chunk: Completely Observed Pair of Strips		
chunkNumber .....	int16	Chunk number
startMu.....	int32	Starting coordinate in great-circle coordinates
endMu.....	int32	Ending coordinate in great-circle coordinates
stripe.....	int32	Stripe number
exportVersion .....	char[32]	Export version
segment; Segment.....	ManyAssoc(sxSegment)	Link to all segments of this chunk
Class Segment: Single Camera Column within a Chunk		
segment_ID; id; segID .....	int64	Unique segment identifier
run.....	int32	Run number of this segment
camCol .....	int32	Camera column number (1–6)
rerun.....	int32	Rerun number
field0.....	int32	Starting field number
nFields .....	int32	Number of fields in segment
dscale.....	int32	Scaling for bias
strip.....	int16	Strip (north or south)
status.....	int16	Status: primary, secondary, or bad
resolved.....	int16	Resolved flag (yes/no)
exported.....	int16	Exported flag (yes/no)
families .....	int16	Families identified flag (yes/no)
loadMJD.....	float64	Date and time at loading of database
equinox.....	float64	Great-circle equinox
node.....	float64	Great-circle ascending node
incl.....	float64	Great-circle inclination
cObs; c_obs .....	float64	CCD clock rate
tracking .....	float64	Tracking rate
raMin; ramin.....	float64	Lowest J2000 right ascension (deg)
raMax; ramax.....	float64	Highest J2000 right ascension (deg)
decMin; decmin.....	float64	Lowest J2000 declination (deg)
decMax; decmax.....	float64	Highest J2000 declination (deg)
lMin; lmin .....	float64	Lowest Galactic longitude (deg)
lMax; lmax .....	float64	Highest Galactic longitude (deg)
bMin; bmin .....	float64	Lowest Galactic latitude (deg)
bMax; bmax .....	float64	Highest Galactic latitude (deg)
lambdaMin; lambdaMin .....	float64	Lowest survey $\lambda$ (deg)
lambdaMax; lambdaMax .....	float64	Highest survey $\lambda$ (deg)
etaMin; etamin.....	float64	Lowest survey $\eta$ (deg)
etaMax; etamax.....	float64	Highest survey $\eta$ (deg)
photoId; photo_id.....	char[32]	Photo ID tag
photoVersion; phot_ver .....	char[32]	Photo version
targetAstroId; tast_id .....	char[32]	Astrometric pipeline used for target
targetAstroVersion; tast_ver.....	char[32]	Astrometric pipeline version used for target selection
exportAstroId; eastr_id .....	char[32]	Astrometric calibration used for export to SX <sup>a</sup>
exportAstroVersion; eastr_ver .....	char[32]	Astrometric pipeline version used for export to SX
targetFCalibId; tfcal_id .....	char[32]	Photometric calibration used for target selection
targetFCalibVersion; tfcal_ver .....	char[32]	Final calibration pipeline version used for target selection
exportFCalibId; efcal_id .....	char[32]	Photometric calibration used for export to SX
exportFCalibVersion; efcal_ver .....	char[32]	Final calibration pipeline version used for export to SX
sxLoaderVersion; loaderVersion; load_ver.....	char[32]	Version of SX Loader used to load this segment
objectSource; obj_src .....	char[32]	Object source
targetSource; targ_src .....	char[32]	Target source
targetVersion .....	char[32]	Target version
skyMask_ ; skyMask .....	uint32[]	Bit-mask-encoded HTM nodes covered by this segment
skyMaskLength_ ; skyMaskLength .....	uint32	Length of bit mask
field; Field.....	ManyAssoc(sxField)	Link to fields in this segment
chunk; Chunk.....	OneAssoc(sxChunk)	Link to the chunk this segment belongs to

<sup>a</sup> The SDSS Science Archive.

TABLE 16  
FIELD CLASS

Name	Type	Description
Class Field: All of the Measured Parameters of a Photometric Field		
fieldID; id; field_ID .....	int64	Unique field identifier: Bit mask consisting of run, rerun, camCol, field
field; fieldNumber .....	int32	Field number
nObjects; nobj .....	int32	Number of objects in the field
nDetectedObjects; ndetected .....	int32	Number of objects actually detected
nChild; nchild .....	int32	Number of “child” objects
nGalaxy; nGals; ngals; ngalaxy; n_gals .....	int32	Number of objects classified as galaxies
nStars; nstars; n_stars .....	int32	Number of objects classified as stars
numStars; numstars; num_stars .....	int32[5]	Number of stars found in each band
nCR; n_CR .....	int32[5]	Number of cosmic rays in each band
nBrightObj; n_bright_obj .....	int32[5]	Number of bright objects in each band
nFaintObj; n_faint_obj .....	int32[5]	Number of faint (nonbright) objects in each band
quality .....	int32	Data quality: AR_DQ_BAD = 1 Not acceptable AR_DQ_ACCEPTABLE = 2 Acceptable but better wanted AR_DQ_GOOD = 3 Fully acceptable AR_DQ_MISSING = 4 Data missing; accepted as hole AR_DQ_HOLE = 5 Data poor; accepted as hole
mjd; MJD .....	float64[5]	MJD (TAI) <sup>a</sup> when row 0 was read
Astrometric Calibration Quantities		
a .....	float64[5]	Astrometric transformation constants
b .....	float64[5]	Astrometric transformation constants
c .....	float64[5]	Astrometric transformation constants
d .....	float64[5]	Astrometric transformation constants
e .....	float64[5]	Astrometric transformation constants
f .....	float64[5]	Astrometric transformation constants
dRow0; d_row0 .....	float64[5]	Zeroth-order row distortion coefficient
dRow1; d_row1 .....	float64[5]	First-order row distortion coefficient
dRow2; d_row2 .....	float64[5]	Second-order row distortion coefficient
dRow3; d_row3 .....	float64[5]	Third-order row distortion coefficient
dCol0; d_col0 .....	float64[5]	Zeroth-order column distortion coefficient
dCol1; d_col1 .....	float64[5]	First-order column distortion coefficient
dCol2; d_col2 .....	float64[5]	Second-order column distortion coefficient
dCol3; d_col3 .....	float64[5]	Third-order column distortion coefficient
csRow; cs_row .....	float64[5]	Slope in row DCR <sup>b</sup> correction for blue objects
csCol; cs_col .....	float64[5]	Slope in column DCR correction for blue objects
ccRow; cc_row .....	float64[5]	Constant row DCR correction for red objects
ccCol; cc_col .....	float64[5]	Constant column DCR correction for red objects
riCut; ri_cut .....	float64[5]	$r^* - i^*$ cutoff between blue and red objects
airmass; air_mass; airMass .....	float64	Air mass for star at frame center (midexposure)
raMin; ra_min .....	float64	Minimum right ascension of field (deg)
raMax; ra_max .....	float64	Maximum right ascension of field (deg)
decMin; dec_min .....	float64	Minimum declination of field (deg)

TABLE 16—Continued

Name	Type	Description
decMax; dec_max.....	float64	Maximum declination of field (deg)
scanlineArea; scanLineArea; scanline_area.....	float32	Area of scan line covered by the field (deg <sup>2</sup> )
stripeArea; stripe_area .....	float32	Area of stripe covered by the field (deg <sup>2</sup> )
Photometric Calibration Quantities		
seqID; seqId.....	int32	Sequence ID overlapping this field (-1 = none)
aa.....	float32[5]	Zero point
aaErr; aa_err.....	float32[5]	Zero-point error
bb.....	float32[5]	Color term
bbErr; bb_err.....	float32[5]	Color term error
cc.....	float32[5]	Color x extinction term
ccErr; cc_err.....	float32[5]	Error in color extinction term
kk.....	float32[5]	Extinction coefficient
kkErr; kk_err.....	float32[5]	Error in extinction coefficient
framesStatus; frames_status.....	int32	Information from the fpFieldStat file: frames processing status
rowOffset; row_offset .....	float32[5]	Offset to add to transformed row coordinates (pixels)
colOffset; col_offset .....	float32[5]	Offset to add to transformed column coordinates (pixels)
saturationLevel; saturation_level .....	int32[5]	Saturation level (counts)
culled.....	int32	Bit mask with the same values as primTarget, Table 27. A set bit indicates that targets of that type were culled from the field, and therefore that field is excluded from that particular science sample
badMask, acceptableMask, goodMask, holeMask.....	int32	Bit masks indicating which quality categories were of the specified quality for each field. The bits are enumerated as follows: AR_FQ_SEEING = 0x1, AR_FQ_TRACKING = 0x2 AR_FQ_PSF = 0x4 AR_FQ_PHOTOMETRY = 0x8 AR_FQ_PHOTO_DROPOUT = 0x10 AR_FQ_ODD_EVEN_DROPOUT = 0x20 AR_FQ_SKY_U = 0x40 AR_FQ_SKY_G = 0x80 AR_FQ_SKY_R = 0x100 AR_FQ_SKY_I = 0x200 AR_FQ_SKY_Z = 0x400 AR_FQ_AMP_U0 = 0x800 AR_FQ_AMP_U1 = 0x1000 AR_FQ_AMP_G0 = 0x2000 AR_FQ_AMP_G1 = 0x4000 AR_FQ_AMP_R0 = 0x8000 AR_FQ_AMP_R1 = 0x10000 AR_FQ_AMP_IO = 0x20000 AR_FQ_AMP_I1 = 0x40000 AR_FQ_AMP_Z0 = 0x80000 AR_FQ_AMP_Z1 = 0x100000 AR_FQ_MANUAL = 0x200000
gain.....	float32[5]	Gain averaged over amplifiers ( $e^- DN^{-1}$ )
darkVariance; dark_variance.....	float32[5]	Dark variance
acceptableMask.....	int32	See description of badMask

TABLE 16—Continued

Name	Type	Description	
goodMask .....	int32	See description of badMask	
holeMask .....	int32	See description of badMask	
Outputs of psp			
skyPsp; sky_psp .....	float32[5]	Sky from the PSF fit ( $\text{asinh mag arcsec}^{-2}$ )	
skyFrames; sky_frames .....	float32[5]	Global sky value in the corrected frame ( $\text{asinh mag arcsec}^{-2}$ )	
skyFramesSub; sky_frames_sub .....	float32[5]	Global sky value after object subtraction ( $\text{asinh mag arcsec}^{-2}$ )	
sky .....	float32[5]	Average sky in frame ( $\text{asinh mag arcsec}^{-2}$ )	
skySig; sky_sig .....	float32[5]	Sigma of sky value distribution ( $\text{asinh mag arcsec}^{-2}$ )	
skyErr; sky_err .....	float32[5]	Error of average sky in frame ( $\text{asinh mag arcsec}^{-2}$ )	
skySlope; sky_slope .....	float32[5]	Slope in sky value along columns ( $\text{asinh mag arcsec}^{-2}$ per field)	
lBias; lBias .....	float32[5]	Left-hand bias level (ADU $\times$ dscale)	
rBias; rBias .....	float32[5]	Right-hand bias level (ADU $\times$ dscale)	
sigPix; sky_pix .....	float32[5]	Clipped $\sigma$ of pixel values in corrected frame ( $\text{asinh mag arcsec}^{-2}$ )	
deVApCorrection; deV_ap_correction .....	float32[5]	De Vaucouleurs aperture correction (mag)	
deVApCorrectionErr; deV_ap_correctionErr .....	float32[5]	De Vaucouleurs aperture correction error (mag)	
expApCorrection; exp_ap_correction .....	float32[5]	Exponential aperture correction (mag)	
expApCorrectionErr; exp_ap_correctionErr .....	float32[5]	Exponential aperture correction error (mag)	
deVModelApCorrection; deV_model_ap_correction .....	float32[5]	De Vaucouleurs model aperture correction (mag)	
deVModelApCorrectionErr; deV_model_ap_correctionErr .....	float32[5]	De Vaucouleurs model aperture correction error (mag)	
expModelApCorrection; exp_model_ap_correction .....	float32[5]	Exponential model aperture correction (mag)	
expModelApCorrectionErr; exp_model_ap_correctionErr .....	float32[5]	Exponential model aperture correction error (mag)	
medianFiberColor; median_fiberColor; median_fiber_color .....	float32[5]	Median fiber colors of objects in field (mag)	
medianPsfColor; median_psfColor; median_psf_color; medianPSFCOLOR .....	float32[5]	Median PSF colors of objects in field (mag)	
q; Q .....	float32[5]	Mean Stokes $Q$ -parameter on the frame	
u; U .....	float32[5]	Mean Stokes $U$ -parameter on the frame	
PSF Fitting Details			
status .....	int32[5]	Type of PSF fit for each filter in the field: AR_PSF_STATUS_UNKNOWN = -1 AR_PSF_STATUS_OK = 0 Second-order parabolas AR_PSF_STATUS_PSF22 = 1 Linear functions AR_PSF_STATUS_PSF11 = 2 Constant AR_PSF_STATUS_NOPSF = 3 Fit from previous frame used AR_PSF_STATUS_ABORTED = 4 No PSF available AR_PSF_STATUS_MISSING = 5 No PSF available	
pspStatus; psp_status .....	int16	Maximum value of “status” over all five filters	
psfNStar; psf_nstar .....	int32[5]	Number of stars used in PSF measurement	
nEffPsf; nEffPSF; neff_psf .....	float32[5]	Effective area of the PSF (pixels)	
psfApCorrectionErr; psf_ap_correctionErr .....	float32[5]	Photometric error due to imperfect PSF model (mag)	
psfSigma1; psf_sigma1 .....	float32[5]	Inner Gaussian $\sigma$ for the composite fit (arcsec)	
psfSigma2; psf_sigma2 .....	float32[5]	Outer Gaussian $\sigma$ for the composite fit (arcsec)	

TABLE 16—Continued

Name	Type	Description
psfB; psf_b .....	float32[5]	Ratio of the amplitude of the inner PSF to the outer PSF at the origin
psfP0; psf_p0.....	float32[5]	The value of the power law at the origin
psfBeta; psf_beta.....	float32[5]	Slope of power law
psfSigmaP; psf_sigmap .....	float32[5]	Width parameter for power law
psfWidth; psf_width .....	float32[5]	Effective PSF width, at the center of the frame (arcsec)
psfPsfCounts; psf_psfCounts.....	float32[5]	Flux via fit to PSF (counts)
psf2GSigma1; psf_sigma1_2G.....	float32[5]	PSF inner $\sigma$ in two-Gaussian fit (arcsec)
psf2GSigma2; psf_sigma2_2G.....	float32[5]	PSF outer $\sigma$ in two-Gaussian fit (arcsec)
psf2GB; psf_b_2G.....	float32[5]	PSF amplitude ratio in two-Gaussian fit (arcsec)
psfCounts; psf_counts .....	float32[5]	PSF counts
nProf; prof_nprof; n_prof.....	int32[5]	Number of profile bins in PSF model
profMeanU; prof_mean_u .....	float32[]	Mean PSF profile, $u$ band
profMeanG; prof_mean_g .....	float32[]	Mean PSF profile, $g$ band
profMeanR; prof_mean_r .....	float32[]	Mean PSF profile, $r$ band
profMeanI; prof_mean_i .....	float32[]	Mean PSF profile, $i$ band
profMeanZ; prof_mean_z .....	float32[]	Mean PSF profile, $z$ band
profMedU; prof_med_u .....	float32[]	Median PSF profile, $u$ band
profMedG; prof_med_g .....	float32[]	Median PSF profile, $g$ band
profMedR; prof_med_r .....	float32[]	Median PSF profile, $r$ band
profMedI; prof_med_i .....	float32[]	Median PSF profile, $i$ band
profMedZ; prof_med_z .....	float32[]	Median PSF profile, $z$ band
profSigU; prof_sig_u .....	float32[]	PSF profile $\sigma$ , $u$ band
profSigG; prof_sig_g .....	float32[]	PSF profile $\sigma$ , $g$ band
profSigR; prof_sig_r .....	float32[]	PSF profile $\sigma$ , $r$ band
profSigI; prof_sig_i .....	float32[]	PSF profile $\sigma$ , $i$ band
profSigZ; prof_sig_z .....	float32[]	PSF profile $\sigma$ , $z$ band
photoobj; PhotoObj; pobj; sxPhotoObj; obj .....	ManyAssoc(PhotoObj)	To-many link to the photometric objects in this field
segment; Segment .....	OneAssoc(Segment)	To-one link to the segment this field is in

<sup>a</sup> International Atomic Time.<sup>b</sup> Differential chromatic refraction.

TABLE 17  
SPECTROSCOPIC PLATE CLASS

Name	Type	Description
Class Plate: Properties of Spectroscopy Plates		
plate_ID .....	int64	Unique ID
mjd; MJD .....	int32	MJD of observation
plateID; plateid .....	int16	Plate ID
tileID; tileid .....	int16	Tile ID
mapID; mapid .....	int16	Map ID
nExp; nexp .....	int16	Number of exposures
tai; TAI .....	float32	TAI seconds
raBoresight; ra_b .....	float64	Right ascension of telescope boresight (deg)
decBoresight; dec_b .....	float64	Declination of telescope boresight (deg)
taiHms; taihms .....	char[20]	Hours, minutes, and seconds
expTime; exptime .....	float32	Total exposure time (s)
expTimeB1; expt_b1 .....	float32	Exposure time, blue camera 1 (s)
expTimeB2; expt_b2 .....	float32	Exposure time, blue camera 2 (s)
expTimeR1; expt_r1 .....	float32	Exposure time, red camera 1 (s)
expTimeR2; expt_r2 .....	float32	Exposure time, red camera 2 (s)
helioRV; heliorv .....	float32	Heliocentric velocity correction ( $\text{km s}^{-1}$ )
ra; radeg .....	float64	Right ascension of plate center (deg)
dec; decdeg .....	float64	Declination of plate center (deg)
htmID .....	int64	HTM level 20 ID of plate center
sn1; SN1 .....	float32[3]	S/N in first spectrograph in $g, r, i$ at 20.2
sn2; SN2 .....	float32[3]	S/N in second spectrograph in $g, r, i$ at 20.2
reddening .....	float32[5]	Reddening in all five bands (mag)
dateObs; dateobs; date_obs .....	char[12]	Date of observation
timeSys; timesys .....	char[8]	Time system
quality .....	char[12]	Quality of night/image
name .....	char[32]	Mapping name: plateID-MJD-sequence
program .....	char[16]	ID of CCD program
version .....	char[20]	Imaging Observer Program version
observer; obsrvr .....	char[20]	Observer name
camVer; camver .....	char[20]	Camera code version
spec2DVer; vers2d; spec2d_ver; spec2d_v...	char[20]	IDLSPEC2D version
utilsVer; versutil; utils_ver; utils_v....	char[20]	IDLUTILS version
spec1DVer; vers1d; vers_1d .....	char[20]	SPEC1D version
readVer; versread .....	char[20]	Version of IDLSPEC2D for preprocessing raw data
combVer; verscomb .....	char[20]	Version of IDLSPEC2D for combining multiple spectra
expID .....	int64[]	Exposure ID array; use <code>expIDString(i)</code> to get the string
expIDString(int32) .....	char*	Function to retrieve exposure ID string
specobj; spec .....	ManyAssocxSpecObj	Link to the spectroscopic objects on this plate

is accessed via

```
objcs/$camCol/
  fpFieldStat-$run-$camCol-$field.fit,
```

and results of the postage-stamp pipeline, which characterizes the point-spread function for a field (§ 4.3), are accessed via

```
objcs/$camCol/
  fpField-$run-$camCol-$field.fit.
```

### 2.5.2. Spectroscopy Files

Spectroscopy files are available from the main Data Archive Server page. Data for individual plates are in sub-directories specified by `$plate/$mjd`, where valid `$plate/` `$mjd` combinations are in Table 4.

Calibrated spectra that have been extracted from the spectroscopic frames with the `spectro2d` pipeline (§ 4.10.1) are named as

```
2d_10/$plate/2dmerge/
  spMerge2d-$mjd-$plate-$fiber.fit,
```

where `$plate` is the plate number (four digits, zero-padded), `$mjd` is the date observations were completed (Table 4), and `$fiber` is the fiber number (three digits, zero-padded). Calibrated spectra, with information from the `spectro1d` pipeline (§ 4.10.2), such as redshift and identification, are named as

```
1d_10/$plate/1d/
  spSpec-$mjd-$plate-$fiber.fit.
```

### 2.6. User Support

We provide help for users at several levels. This paper is a general introduction to the data products. Our Web sites have additional documentation. We maintain a set of frequently asked questions with answers to help people become more familiar with how to use the data set. There are also two mailing lists that we encourage users to join. The first is an *SDSS Early Data Release Users' Group*, which we will use to communicate periodically with users. The second is the *SDSS Early Data Release Users' Forum*, where we encourage all users to post questions and suggestions. We

TABLE 18  
CONSTANTS USED IN THE SDSS

Name	Type	Description
Class SDSSConstants: Definition of the Survey's Geometry		
siteLongitude .....	float64	Site geodesic west longitude (deg)
siteLatitude .....	float64	Site geodesic north latitude (deg)
siteAltitude .....	float64	Altitude (m)
surveyCenterRa.....	float32	J2000 right ascension of survey center (deg)
surveyCenterDec.....	float32	J2000 declination of survey center (deg)
surveyEquinox .....	float32	Survey equinox (yr)
lambdaMin .....	float32	Survey longitude minimum (deg)
lambdaMax .....	float32	Survey longitude maximum (deg)
etaMin.....	float32	Survey latitude minimum (deg)
etaMax.....	float32	Survey latitude maximum (deg)
lambdaMinNStripe.....	float32[45]	Minimum longitude of each northern stripe
lambdaMaxNStripe.....	float32[45]	Maximum longitude of each northern stripe
etaNStripe .....	float32[45]	Latitude for each northern stripe (deg)
southStripeNumbers.....	int32[3]	The southern stripe numbers
lambdaMinSStripe.....	float32[3]	Minimum longitude of each southern stripe
lambdaMaxSStripe.....	float32[3]	Maximum longitude of each southern stripe
etaSStripe .....	float32[3]	Latitude for each southern stripe
stripeWidth .....	float32	Width of stripes (deg)
scanSeparation.....	float64	Scan separation (deg)
stripeSeparation.....	float32	Stripe separation (deg)
brickLength .....	float32	Length of imaging bricks (deg)
ccdColSep .....	float64	CCD column separation (deg)
northMajor .....	float32	Major-axis radius of northern survey (deg)
northMinor .....	float32	Minor-axis radius of northern survey (deg)
northPA.....	float32	Position angle of northern survey area (deg)
telescopeSize .....	float32	Nominal diameter of primary mirror (m)
fieldOfView .....	float32	Telescope field of view (deg)
chipPixelX .....	int32	Chip X-size (pixels)
chipPixelY .....	int32	Chip Y-size (pixels)
nFilters .....	int32	Number of filters
nChipsPerFilter.....	int32	Number of chips per filter
effectiveLambda.....	int32[5]	Effective wavelengths of each filter
limitingMagnitudes.....	float32[5]	Limiting magnitudes per band
filters .....	char[16]	Filter names
refBand .....	char[3]	Reference band
specPerPlate .....	int32	Number of spectra taken per plate
specLowerLimit.....	int32	Spectrograph wavelength lower limit ( $\text{\AA}$ )
specUpperLimit.....	int32	Spectrograph wavelength upper limit ( $\text{\AA}$ )
resolutionPLDL; resolution...	int32	Resolution of spectrograph ( $\lambda/\Delta\lambda$ )

will monitor this mailing list and post answers, or pointers to on-line documentation, as appropriate. We also support a help desk, which is accessible via E-mail and telephone, as detailed on our Web sites.

### 2.7. Acknowledging SDSS Data

Access to the SDSS EDR is open. These data may be used for any astronomical research. Data from the SDSS public archive may not be used for any commercial publication or other commercial purpose except with explicit approval by the Astrophysical Research Consortium.

Noncommercial and technical publications based on the SDSS EDR should contain the following acknowledgment:

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 US Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS Web site is <http://www.sdss.org/>.

The Participating Institutions are the University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, the Johns Hopkins University, the Max Planck Institute for Astronomy (MPIA), the Max Planck Institute for Astrophysics (MPA), New Mexico State University, Princeton University, the United States Naval Observatory, and the University of Washington.

## 3. HARDWARE

### 3.1. 2.5 m Telescope and Instruments

The 2.5 m telescope and its somewhat unusual optical design (Waddell et al. 1998) are described in detail in York et al. (2000) and the SDSS Project Book, available on the Web.<sup>64</sup> The camera and the pair of double spectrographs are discussed there as well (see also Gunn et al. 1998). We

<sup>64</sup> See <http://astro.princeton.edu/PBOOK/welcome.htm>.

review here only details that are of direct relevance to working with the data.

### 3.2. Imaging System

The parameters of direct relevance to the imaging system of the SDSS are presented in Table 1. Since York et al. (2000) was published, we have learned a great deal about our filter and detector system, and a preliminary discussion of these matters is given below, as are some details concerning the geometry of the great-circle time-delay integration (TDI, or drift-scan) stripes that cover the survey area.

#### 3.2.1. Filters

The situation with the SDSS response functions for the five filter passbands, and the resulting photometric system, is rather complex. There is a set of primary standard stars that have been measured at the US Naval Observatory (USNO) 40 inch (1 m) telescope and with the SDSS PT (§ 3.4), which together define a photometric system that we believe to be self-consistent to approximately 1%; this system is roughly as described in Fukugita et al. (1996). These primary standards are described further in § 4.5 below. Unfortunately, the filters used on the USNO telescope and the PT differ systematically from those on the 2.5 m camera, and we still do not have a complete understanding of the transformations between these two systems.<sup>65</sup> Thus, the photometric system defined by the USNO telescope and the PT is not directly applicable to the 2.5 m data, as described in detail in § 4.5. Figure 4 gives the average measured quantum efficiencies of the 2.5 m camera detectors multiplied by the reflectivity of the primary and secondary (the two transmissive surfaces have negligible effect on the throughput); curves are given both assuming no atmosphere and including the transmission of the atmosphere above Apache Point on a night of average humidity at air mass 1.3. Tables containing the system response in each filter are available on our Web sites. The thinned CCDs also suffer from internal scattering that scatters light longward of roughly 6000 Å into an extended halo around an object; this decreases the effective quantum efficiency for a point source. For extended sources ( $>30''$ ), this effect is negligible, and the dashed curves indicate the quantum efficiency in this case in the *r* and *i* filters. The *z* chip is thick and does not suffer from this problem.

The camera responses were measured by an instrument with a roughly triangular wavelength response with full width at half-maximum (FWHM) about 100 Å; this resolution has *not* been corrected for in these data but does not appreciably alter the shapes. Better and more detailed response data will be obtained and published later, but the results here are adequate for most purposes.

Table 19 gives corresponding properties of the filters, updating those tabulated in Fukugita et al. (1996) and Fan et al. (2001b): the effective wavelength of each filter  $\lambda_{\text{eff}}$ , the photon-weighted mean of the quantity  $\ln^2(\lambda/\lambda_{\text{eff}})$  (a measure of the effective width of the filter), the FWHM of the fil-

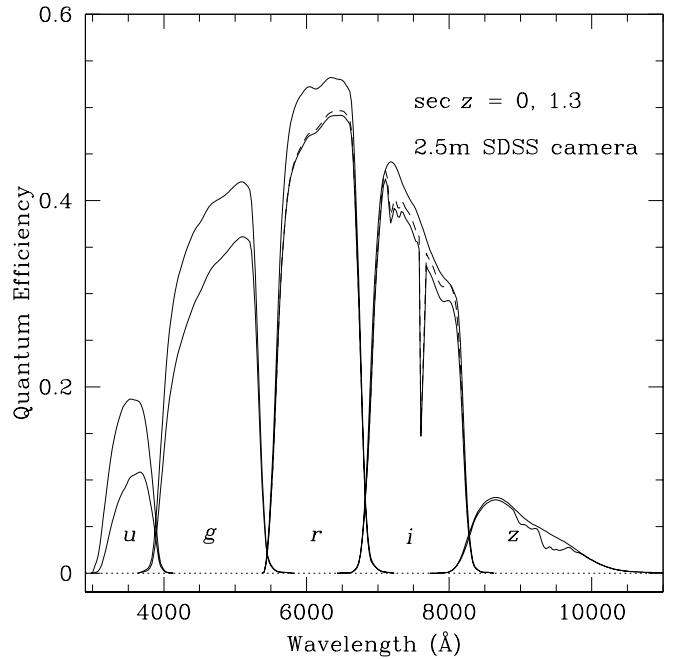


FIG. 4.—Preliminary 2.5 m telescope filter responses, in *u*, *g*, *r*, *i*, and *z*. The upper curve in each case is the filter response including the quantum efficiency of the CCD and the reflectivity of the primary and secondary, ignoring the atmosphere, and the lower curve assumes an air mass of 1.3. Scattering within the thin chips affects the *r* and *i* bands; this has no effect on extended objects, and the corresponding response curve is given by the dashed line in these two cases.

ter, and  $Q$ , the integral of the system efficiency over  $d \ln \lambda$ , effective quantum efficiency (all assuming 1.3 air masses, and observing a point source). This last quantity relates the measured apparent magnitude (on an AB system) to the number of detected electrons:

$$\text{No. } e^- = (1.96 \times 10^{11} t Q) \times 10^{-0.4m}, \quad (1)$$

where  $t$  is the exposure time in seconds.

As pointed out in § 1, we refer to data on the standards system with the magnitude labels  $u'g'r'i'z'$ , and the provisional 2.5 m magnitudes with the labels  $u^*g^*r^*i^*z^*$ . The SDSS photometry itself is presented in the provisional 2.5 m system. Finally, the 2.5 m filters themselves are referred to in this paper simply as *u*, *g*, *r*, *i*, and *z*, a change from some of our earlier papers.

#### 3.2.2. Great-Circle Drift-scanning

The survey coordinate system  $(\lambda, \eta)$  is a spherical coordinate system with poles at  $\alpha = 95^\circ$ ,  $\delta = 0^\circ$  and  $\alpha = 275^\circ$ ,  $\delta = 0^\circ$  (J2000). The survey equator is thus a great circle per-

TABLE 19  
SDSS FILTER PARAMETERS

Name	$\lambda_{\text{eff}}$ (Å)	$\sigma^2$	FWHM (Å)	$Q$
<i>u</i> .....	3551	$3.00 \times 10^{-3}$	581	0.0171
<i>g</i> .....	4686	$7.13 \times 10^{-3}$	1262	0.0893
<i>r</i> .....	6166	$3.13 \times 10^{-3}$	1149	0.0886
<i>i</i> .....	7480	$2.58 \times 10^{-3}$	1237	0.0591
<i>z</i> .....	8932	$3.18 \times 10^{-3}$	994	0.0099

<sup>65</sup> In fact, strictly speaking, there is not one 2.5 m system but six, because the detectors and filters in each CCD column are slightly different. These differences are appreciable only in the *z* band, which is cut off on the long-wavelength side by the detector response. While these differences must be accounted for in detail for some work, for most purposes the average response curves are sufficient.

pendicular to the J2000 celestial equator, intersecting it at  $\alpha = 185^\circ$  and  $\alpha = 5^\circ$ . Lines of constant  $\eta$  are great circles perpendicular to the survey equator, and lines of constant  $\lambda$  are small circles parallel to the survey equator;  $\lambda = 0^\circ$ ,  $\eta = 0^\circ$  is located at  $\alpha = 185^\circ$ ,  $\delta = 32^\circ 5'$ , with  $\eta$  increasing northward.

The survey area is divided into stripes, where each stripe is centered along a line of constant  $\eta$ , separated from the adjoining stripe or stripes by  $2^\circ 5'$ . Each drift scan tracks a survey stripe, offset by  $\pm 386''$  perpendicular to the stripe. Two scans (or “stripes”), one offset to the north and one to the south, are required to fill a stripe. The survey latitude tracked by stripe  $n$  is given by

$$\eta = (n - 10) \times 2^\circ 5' - 32^\circ 5' \quad (2)$$

in the northern Galactic hemisphere, and by

$$\eta = (n - 82) \times 2^\circ 5' - 32^\circ 5' \quad (3)$$

for the three stripes in the southern Galactic hemisphere. These stripes are superposed on a Galactic extinction map in Figure 2 of York et al. (2000).

The natural coordinate system to use for processing a given drift scan is the *great-circle coordinate system* for that stripe,  $(\mu, \nu)$ , in which the equator of the coordinate system is the great circle tracked by the scan. This great circle is inclined by  $i = \eta + 32^\circ 5'$  to the J2000 celestial equator, with an ascending node of  $95^\circ$ ;  $\mu = \alpha$  at the ascending node, and  $\mu$  increases in the scan direction (east) and  $\nu$  increases to the north. Each stripe has its own great-circle coordinate system.

For reference, the equations to transform among the different coordinate systems are

$$\begin{aligned} \cos(\alpha - 95^\circ) \cos \delta &= -\sin \lambda \\ &= \cos(\mu - 95^\circ) \cos \nu, \\ \sin(\alpha - 95^\circ) \cos \delta &= \cos \lambda \cos(\eta + 32^\circ 5') \\ &= \sin(\mu - 95^\circ) \cos \nu \cos i - \sin \nu \sin i, \\ \sin \delta &= \cos \lambda \sin(\eta + 32^\circ 5') \\ &= \sin \nu \cos i + \sin(\mu - 95^\circ) \cos \nu \sin i. \end{aligned}$$

### 3.3. Spectroscopic System

We produce 640 individual spectra in a  $3^\circ$ -diameter field at a resolution  $R \equiv \lambda/\Delta\lambda$  of about 1800 in the wavelength range of 3800–9200 Å. This wavelength range is divided between two cameras by a dichroic at about 6150 Å, and there are two spectrographs, each producing 320 spectra. There are thus four CCD detectors, each of the same kind as are present in the  $g$ ,  $r$ , and  $i$  bands in the camera, 2048 pixels square with  $24 \mu\text{m}$  pixels. The spectroscopic system is discussed in York et al. (2000). Results from commissioning the system are discussed in Castander et al. (2001).

The fibers carrying the light from the drilled plug plates to the spectrographs subtend about  $3''$  in the focal plane, and they are imaged in turn in the spectrograph cameras with a footprint of about 3 pixels. The straight-through transmissive immersion grisms produce a dispersion that is roughly linear in logarithmic wavelength. The spectrographs are very efficient; quantum efficiencies on the sky as measured from standard stars as a function of wavelength for each of the four spectrographic CCDs are presented in Figure 5.

They peak at over 25% in the red, and at just under 20% in the blue.

The nominal exposure time for each plate is 45 minutes, split into at least three parts for cosmic-ray rejection, with the exact number determined by observing conditions. This set of science exposures is preceded and followed by a series of shorter exposures for calibration: arcs, flat fields, and a 4 minute *smear* exposure on the sky for spectrophotometric calibration, in which the telescope is moved so that the  $3''$  fiber on each object effectively covers a  $5'' \times 8''$  aperture, aligned with the parallactic angle. The smear exposures allow us to account for object light excluded from the  $3''$  fibers because of seeing and atmospheric refraction; they provide an accurate (albeit low signal-to-noise ratio) measure of the true spectral shape of the objects and are used for spectrophotometric calibration.<sup>66</sup> The calibration and science exposures are immediately processed through a streamlined version of the second spectroscopic pipeline (§ 4.10) to inform the observers whether the calibrations were successful and to provide signal-to-noise ratio (S/N) diagnostics on the science exposures.

For each science exposure, the  $(S/N)^2$  per pixel through the SDSS imaging passbands is measured and evaluated as a function of fiber magnitude for each spectrograph camera. We take repeated 15 minute exposures until the cumulative median  $(S/N)^2 > 15$  at  $g^* = 20.2$  and  $i^* = 19.9$  in all four cameras. In clear, non-Moony conditions, the  $(S/N)^2$  threshold is easily reached in three exposures, and we never take fewer than three; in (partially) cloudy or Moony conditions, more exposures may be required.

### 3.4. Photometric Telescope

We use the 20 inch PT, located next to the 2.5 m telescope’s enclosure, to measure nightly extinctions and to observe transfer fields (secondary patches) that in turn are used to calibrate the 2.5 m imaging data. Details of the full photometric calibration process can be found in § 4.5. This telescope is a commercial reflector built by DFM Engineering (Longmont, Colorado), modified to incorporate improved baffling and a coma corrector to increase its field of view. It is equipped with a thinned SITe 2048 × 2048 chip with  $24 \mu\text{m}$  square pixels like the  $u$  chips in the 2.5 m camera, and a set of filters nominally the same as those in the camera (see § 3.2.1).

The telescope operates automatically, observing primary standard stars and secondary-patch transfer fields selected from an on-line database. Observing staff can monitor progress with real-time tools that display cloud cover, extinction coefficients, and observing progress (Hogg et al. 2001).

### 3.5. Data Acquisition

The data acquisition system (Petravick et al. 1994) records information from the imaging camera, spectrographs, and photometric telescope. Data are transferred via magnetic tape, with critical, low-volume samples sent over the Internet. Each system uses report files to track the observations.

Data from the imaging camera are collected in the TDI mode. We treat the data from each imaging camera column

<sup>66</sup> Note that we only implemented the smear exposures halfway through commissioning. Table 4 indicates for each plate whether a smear exposure was actually taken.

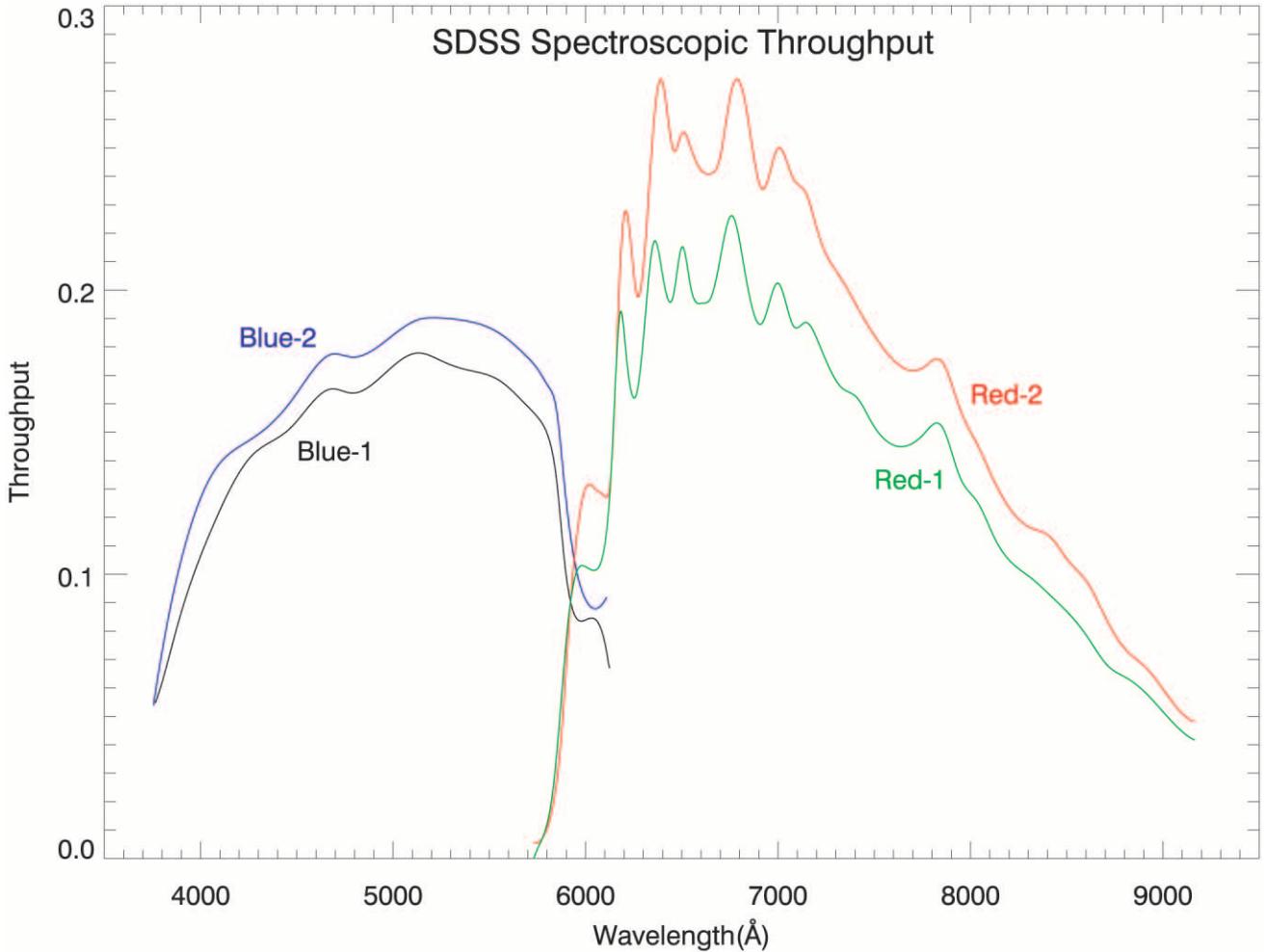


FIG. 5.—Quantum efficiency of atmosphere-telescope-fiber-spectrograph combination for each of the four spectroscopic chips, as measured from F stars with calibrated photometry.

of five photometric and two astrometric CCDs as a *scan line*. For convenience, data from each CCD are broken into *frames* containing 1361 lines. Before processing, the 128 rows from the next frame are added to the top of each frame, so that the pipelines work on  $2048 \times 1489$  images. The resulting overlap between reduced frames (128 rows) is roughly the same as the number of columns that overlap with the other strip of a stripe. Some objects are detected in more than one frame, but when loading the databases we mark one of these detections as the “primary” detection (see the discussion in § 4.7). The frames that correspond to the same sky location in each of the five filters are grouped together for processing as a *field*. Frames from the astrometric CCDs are not saved, but rather, stars from them are detected and measured in real time to provide feedback on telescope tracking and focus. These measurements are also written to magnetic tape. This same analysis is done for the photometric CCDs, and we save these results along with the actual frames. Each night, a special bias run is taken to monitor the bias levels on the CCD amplifiers.

Data from the spectrographs are read from the four CCDs (one red channel and one blue channel in each of the two spectrographs) after each exposure. A complete set of exposures includes bias, flat, arc, smear, and science exposures taken through the fibers, as well as a uniformly illuminated flat to take out pixel-to-pixel variations.

Data from the PT include bias frames, dome and twilight flats for each filter, measurements of primary standards in each filter, and measurements of our secondary calibration patches in each filter.

All of these systems are supported by a common set of observers’ programs, with observer interfaces customized for each system to optimize our observing efficiency.

## 4. SOFTWARE

### 4.1. Data Processing Factory

Data from APO are transferred to Fermilab for processing and calibration. Three flavors of data are produced at APO: data from the imaging camera, data from the photometric telescope, and spectra from the spectrographs. Imaging data are processed with the imaging pipelines: the *astrometric pipeline* (*astrom*; § 4.2) performs the astrometric calibration; the *postage-stamp pipeline* (*psp*; § 4.3) characterizes the behavior of the point-spread function as a function of time and location in the focal plane; the *frames pipeline* (*frames*; § 4.4) finds, deblends, and measures the properties of objects; and the *final calibration pipeline* (*nfcalib*; § 4.5.3) applies the photometric calibration to the objects. This calibration uses the results of the PT data processed with the *monitor telescope pipeline* (*mtpipe*; § 4.5.2).

TABLE 20  
PIPELINE VERSIONS USED FOR EARLY DATA RELEASE

Run	astrom <sup>a</sup>	psp <sup>b</sup>	frames <sup>c</sup>	mtpipe <sup>d</sup>	nfcalib <sup>e</sup>	resolve <sup>f</sup>	target <sup>g</sup>	plate <sup>h</sup>	spectro2d <sup>i</sup>	spectro1d <sup>j</sup>
94.....	v3.3	v5.2.21	v5.2.21	v7.2	v1.7.8	v2.7	v2.7	v2.6	v4.6.2	v5.3.2
125.....	v3.3	v5.2.21	v5.2.21	v7.2	v1.7.8	v2.7	v2.7	v2.6	v4.6.2	v5.3.2
752.....	v3.3	v5.2.6	v5.2.6	v7.2	v1.7.8	v2.2a	v2.2a	v2.3	v4.6.2	v5.3.2
756.....	v3.2	v5.2.6	v5.2.6	v7.2	v1.7.8	v2.2a	v2.2a	v2.3	v4.6.2	v5.3.2
1336....	v3.3	v5.2.12	v5.2.12	v7.2	v1.7.8	v2.5	v2.5	v2.6	v4.6.2	v5.3.2
1339...	v3.3	v5.2.12	v5.2.12	v7.2	v1.7.8	v2.5	v2.5	v2.6	v4.6.2	v5.3.2
1356...	v3.3	v5.2.12	v5.2.12	v7.2	v1.7.8	v2.5	v2.5	v2.6	v4.6.2	v5.3.2
1359...	v3.3	v5.2.10	v5.2.10	v7.2	v1.7.8	v2.5	v2.5	v2.6	v4.6.2	v5.3.2

<sup>a</sup> Provides the astrometric calibration for imaging data.

<sup>b</sup> The postage-stamp pipeline (psp) characterizes the PSF and sky level during an imaging run.

<sup>c</sup> Finds and measures objects in imaging data.

<sup>d</sup> Finds and measures objects in data from the photometric telescope.

<sup>e</sup> Final photometric calibration; nfcalib matches objects from frames and mtpipe to calibrate imaging data.

<sup>f</sup> Chooses among multiple detections of objects in overlap areas to define the primary set of detected objects.

<sup>g</sup> Selects primary objects for spectroscopic observation.

<sup>h</sup> Defines locations of holes in an aluminum plug plate for spectroscopic targets.

<sup>i</sup> Extracts and flux-calibrates spectra from the full spectroscopic frames.

<sup>j</sup> Identifies and measures the extracted spectra.

The combination of the psp and frames pipelines is sometimes referred to as “Photo.”

Individual imaging runs that interleave are prepared for spectroscopy with the following steps: *resolve* (§ 4.7) selects a primary detection for objects that fall in an overlap area; the *target selection pipeline* (target; § 4.8) selects objects for spectroscopic observation; and the *plate pipeline* (plate; § 4.9) specifies the locations of the plates on the sky and the location of holes to be drilled in each plate. Spectroscopic data are first extracted and calibrated with the *two-dimensional pipeline* (spectro2d; § 4.10.1) and then classified and measured with the *one-dimensional pipeline* (spectro1d; § 4.10.2).

The EDR was prepared using the versions of pipelines indicated in Table 20. The data for the EDR were reduced with a consistent set of pipeline versions, with only minor version changes to address operational issues.

We continue to develop these pipelines, and in what follows, we describe known problems and future developments where relevant. We also continue to improve photometric calibration techniques (§ 4.5).

#### 4.2. Astrometric Pipeline

A separate great-circle coordinate system is defined for each stripe (§ 3.2.2). In these systems, the stripe center is the equivalent of the equator in the equatorial ( $\alpha, \delta$ ) system. Pixel coordinates are corrected for empirically derived optical distortion terms, and the resulting mapping from corrected CCD row and column pixel positions to great-circle longitude and latitude is linear to a very good approximation. Astrometric solutions are carried out in this coordinate system. One of two reduction strategies is employed, depending on the coverage of astrometric catalogs:

1. Whenever possible, stars detected on the  $r$  photometric CCDs are matched directly with stars in the USNO CCD Astrograph Catalog (UCAC; Zacharias et al. 2000). The UCAC extends down to  $R = 16$ , giving approximately 2 mag of overlap with unsaturated stars on the photometric  $r$  CCDs. For the EDR, this technique was used for the south equatorial stripe (runs 94 and 125) and for the north equatorial stripe (runs 752 and 756).

2. If UCAC coverage is not available to reduce an imaging run, bright stars detected with the astrometric CCDs are matched with the Tycho-2 Catalogue (Høg et al. 2000). The fainter stars on the astrometric CCDs are then matched with the brighter detections on the  $r$  CCDs, enabling us to transfer the pixel coordinates of the catalog stars onto the  $r$  CCD coordinate system. In the EDR, this technique was used for the SIRTF area—runs 1336, 1339, 1356, and 1359.

For each  $r$  frame, these mappings result in an affine transformation relating corrected pixel positions to celestial coordinates. A secondary catalog is produced from the detections on the  $r$  CCDs. This secondary catalog is then matched to centroid positions on the  $i$ ,  $u$ ,  $z$ , and  $g$  CCDs to derive affine transformations in those filters. The transformation also includes terms to correct for differential chromatic refraction, and those terms are applied when the colors of objects are known (Table 16). Positions of detected objects given in the EDR have had this correction applied.

##### 4.2.1. Astrometric Quality

The relative astrometry between the  $r$  and the  $i$ ,  $u$ ,  $z$ , and  $g$  CCDs is independent of the astrometric catalog used and typically has rms errors of 20–30 mas per coordinate, and systematic errors on the order of 20 mas. The quality of the absolute astrometry (based on the  $r$  astrometric solutions and centroids) is dependent on the astrometric catalog used and is dominated by systematic errors that vary on time-scales of minutes. Within a given run, the distribution of systematic errors is well characterized by a Gaussian. Reductions against the UCAC have rms systematic errors on the order of 50 mas per coordinate. Reductions against Tycho-2 have rms systematic errors on the order of 100 mas and show additional systematic errors constant over entire scans of up to 50 mas. Centroiding errors contribute an additional random source of error, on the order of 20 mas, for objects brighter than  $r^* = 20$ . Comparison with the astrometry of the Two Micron All Sky Survey (2MASS) for stars in common shows systematic offsets under 50 mas, well within our quoted errors and the expected systematic astrometric calibration effects quoted by the 2MASS team.

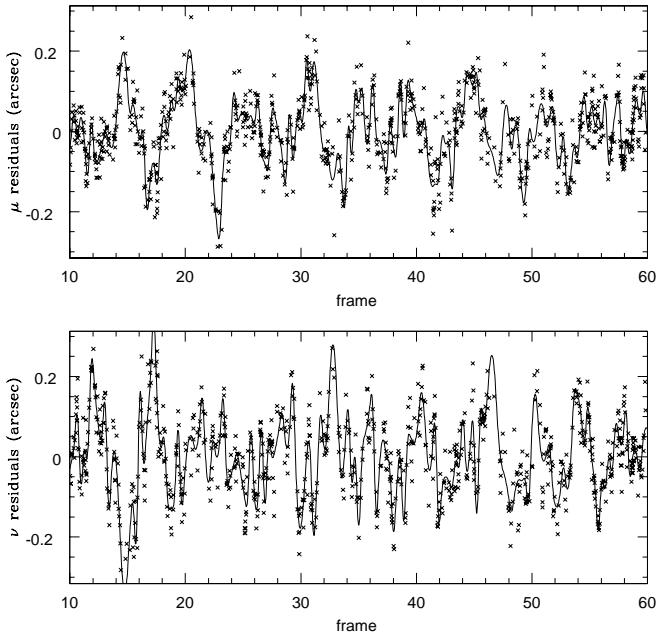


FIG. 6.—Typical astrometric residuals, comparing positions between the astrometric and photometric chips for objects in common (run 1356, CCD 73 vs. 13). The line going through the points is a spline fit to these residuals, attempting to remove atmospheric fluctuations.

Atmospheric conditions contribute significantly to image wander. These affect the Tycho-2 reductions more than the UCAC reductions because of the shorter integration times on the astrometric CCDs. We attempt to follow this wander by fitting the residuals with cubic splines. Figure 6 shows the astrometric residuals as a function of frame number for a typical run. The spline-fitted solution is superposed on the points showing the residuals. The top panel shows  $\mu$ -residuals (along the direction of the scan), and the bottom panel shows  $\nu$ -residuals (the cross-scan direction). Approximately 100 frames are obtained per hour, so the figure shows roughly half an hour of scanning. Note that the residuals wander several tenths of an arcsecond over timescales of minutes. The frequency and amplitude of these wanderings vary from night to night (and, occasionally, hour to hour).

#### 4.2.2. WCS

The bias-subtracted, flat-fielded data frames in the EDR include world coordinate system information in the FITS file headers. This information enables some display and analysis software to provide right ascension and declination information on a pixel-by-pixel basis and also to overlay an equatorial coordinate grid on the image.

Present WCS proposed standards do not fully support a rigorous transformation from great-circle to equatorial coordinates. As a result, the WCS representation does not reflect the full accuracy of the astrometric solution, but the accuracy is better than 1 pixel (about  $0.^{\prime\prime}4$ ) within a frame.

The conversion from row and column measured in a field to  $(\alpha, \delta)$  (J2000 degrees) is

$$\alpha = \text{CRVAL2} + (\Delta_{\text{COL}} \text{CD2\_1} + \Delta_{\text{ROW}} \text{CD2\_2}) / \cos(\text{CRVAL1}), \quad (4)$$

$$\delta = \text{CRVAL1} + \Delta_{\text{COL}} \text{CD1\_1} + \Delta_{\text{ROW}} \text{CD1\_2}, \quad (5)$$

where  $\Delta_{\text{COL}} = \text{column} - \text{CRPIX1}$  and  $\Delta_{\text{ROW}} = \text{row} - \text{CRPIX2}$  and CRVAL1, CRVAL2, CRPIX1, CRPIX2, CD1\_1, CD1\_2, CD2\_2, and CD2\_1 are parameters defined in the FITS header for each field.

#### 4.3. Postage-Stamp Pipeline

As mentioned above, the data stream from each CCD is divided into an overlapping series of  $10' \times 13.^{\circ}5$  frames, for ease of processing; the frames pipeline (§ 4.4) will process these separately. However, in order to ensure continuity along the data stream, certain quantities need to be determined on timescales up to the length of the imaging run. The astrometric and photometric calibrations certainly fall into that category; in addition, the postage-stamp pipeline (psp) calculates a global sky for a field, flat-field vector, bias level, and the point-spread function (PSF).

Even in the absence of atmospheric inhomogeneities, the SDSS telescope delivers images whose FWHMs vary by up to 15% from one side of a CCD to the other; the worst effects are seen in the chips farthest from the optical axis. Moreover, since the atmospheric seeing varies with time, the delivered image quality is a complex two-dimensional function even on the scale of a single frame. An example of the instantaneous image quality across the imaging camera is shown in Figure 7, where each rectangle represents one chip.<sup>67</sup>

The description of the point-spread function is critical for accurate PSF photometry, for star-galaxy separation, and for unbiased measures of the shapes of nonstellar objects; we need to map the full variation of the PSF even on scales of a single frame. The SDSS imaging PSF is modeled heuristically in each band using a Karhunen-Loëve (K-L) transform (Lupton et al. 2001). In particular, using stars brighter than roughly 20th magnitude, we expand the PSF from a series of five frames into eigenimages and keep the first three terms. We fit the variation of the coefficients by multiplying these terms to second order in position across the chip, using data from the frame in question, plus the immediately preceding and following half-frames.

The success of this K-L expansion is gauged by comparing PSF photometry based on the modeled K-L PSFs with large-aperture photometry for the same (bright) stars. The width of the distribution of these differences is typically 1% or less, which is thus an upper limit on the accuracy of the PSF photometry (not including calibration problems; see § 4.5). Without accounting for the spatial variation of the PSF across the image, the photometric errors would be as high as 15%. We have recently found a subtle dependence of the PSF width on stellar color in the  $g$  band; this affects PSF photometry at less than the 2% level and will be addressed in future data releases.

Parameters that characterize one frame of imaging data are stored in the class Field (Table 16). The status parameter flag for each frame indicates the success of the K-L decomposition; its possible values are given in Table 16. In particular, if the data do not support the fitting of a second-order term to the variation of the coefficients with position, a linear fit is carried out and status is set to 1. If even this is not warranted by the data, the coefficients are set to be constants, and the status flag is set to 2. Finally, if no PSF stars

<sup>67</sup> The situation has improved since these data were taken, as the telescope is now better collimated.

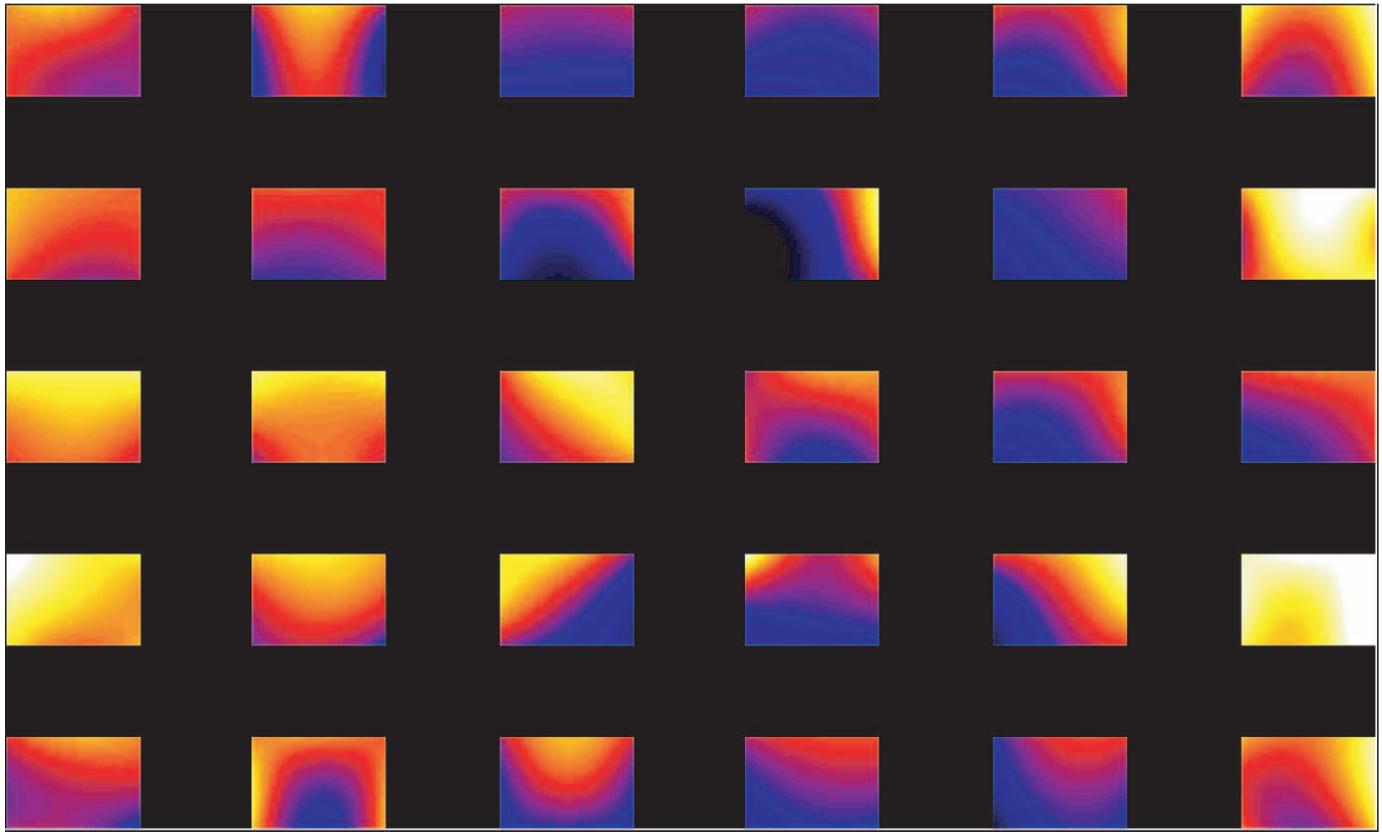


FIG. 7.—Image quality across the imaging camera. The width of the PSF is a strong function of position in the camera, even on the scale of a single CCD. Each square represents one of the photometric CCDs. The colors indicate the PSF width across the camera at a single moment of relatively poor seeing; the variations are due to time variability during the 54 s integration, plus effects in the telescope optics. The stretch is linear, from  $1''4$  to  $2''$ .

are available at all, the PSF model is set to that of the previous frame and `status = 3`. A more quantitative measure of the accuracy of the PSF fit on a given frame is given by the scatter in the difference between PSF magnitudes and aperture magnitudes, as reported in `psfApCorrectionErr`. Note that the actual K-L values and the eigenshapes are not reported in the tables, so the shape of the PSF as a function of position within a CCD cannot be reconstructed based on the parameters included in the EDR.

The postage-stamp pipeline returns various measures of the PSF shape in addition to the K-L decomposition, including parameters of the best-fit double Gaussian, evaluated at the center of each frame. These are the `psf2GSigma` parameters in the `Field` class. The `psfWidth` is the effective width, also determined at the center of each frame. It is a good generic number to quote for the seeing on each frame. Figure 8 shows the `psfWidth` in *r* for each CCD column in each run of the EDR imaging data. Improvements in telescope collimation and thermal environment since these data were taken have given rise to substantially better seeing.

The pipeline calculates a PSF aperture correction for each field. We did not fully test this correction and did not properly apply it to the EDR. The quantity `apCorrRunErr` is the median value of the difference between `psfMag` and the aperture magnitude measured with a  $7''.43$  radius aperture over the bright stars in the frame and is a measure of the limitations of our K-L decomposition of the PSF. This quantity is Gaussian-distributed with  $\sigma \approx 0.03$  mag, but `apCorrRunErr` can be as large as 0.1 mag in regions where the PSF is changing rapidly (FWHM changing by more than 10% on a

single frame); adding it directly to `psfMag` in the EDR improves the overall PSF photometry accordingly. This correction will be applied to the data correctly in future releases of the SDSS data. The value of `apCorrRunErr` is shown for each run in Figure 9, and we provide a table of these corrections on our Web sites.

#### 4.4. Frames Pipeline

The frames pipeline (`frames`) detects, deblends, and measures objects, carrying out this processing on a field-by-field basis. This section describes what one needs to know to interpret the quantities we calculate in `frames`, rather than the technical details, which will be covered in future papers. These quantities are stored in the `PhotoObj` and `Profile` classes (Table 6). Mask bits set in `objFlags` of `PhotoObj` are for the object detections for all bands combined. Mask bits set in `flags` of `PhotoObj` are for the detection in each band. The mask-bit values are given in Table 9. We refer to these classes and masks throughout this section. Further products of the imaging pipelines are given in § 2.2.

##### 4.4.1. How frames Designates Objects

Each object detected during the `frames` analysis of a particular set of data is given a unique identifier, which consists of five integers:

run of class `Segment` (Table 15): TA run number refers to a given uninterrupted drift scan.

rerun of class `Segment` (Table 15): Each time we analyze a run with a different set of pipeline versions or calibrations,

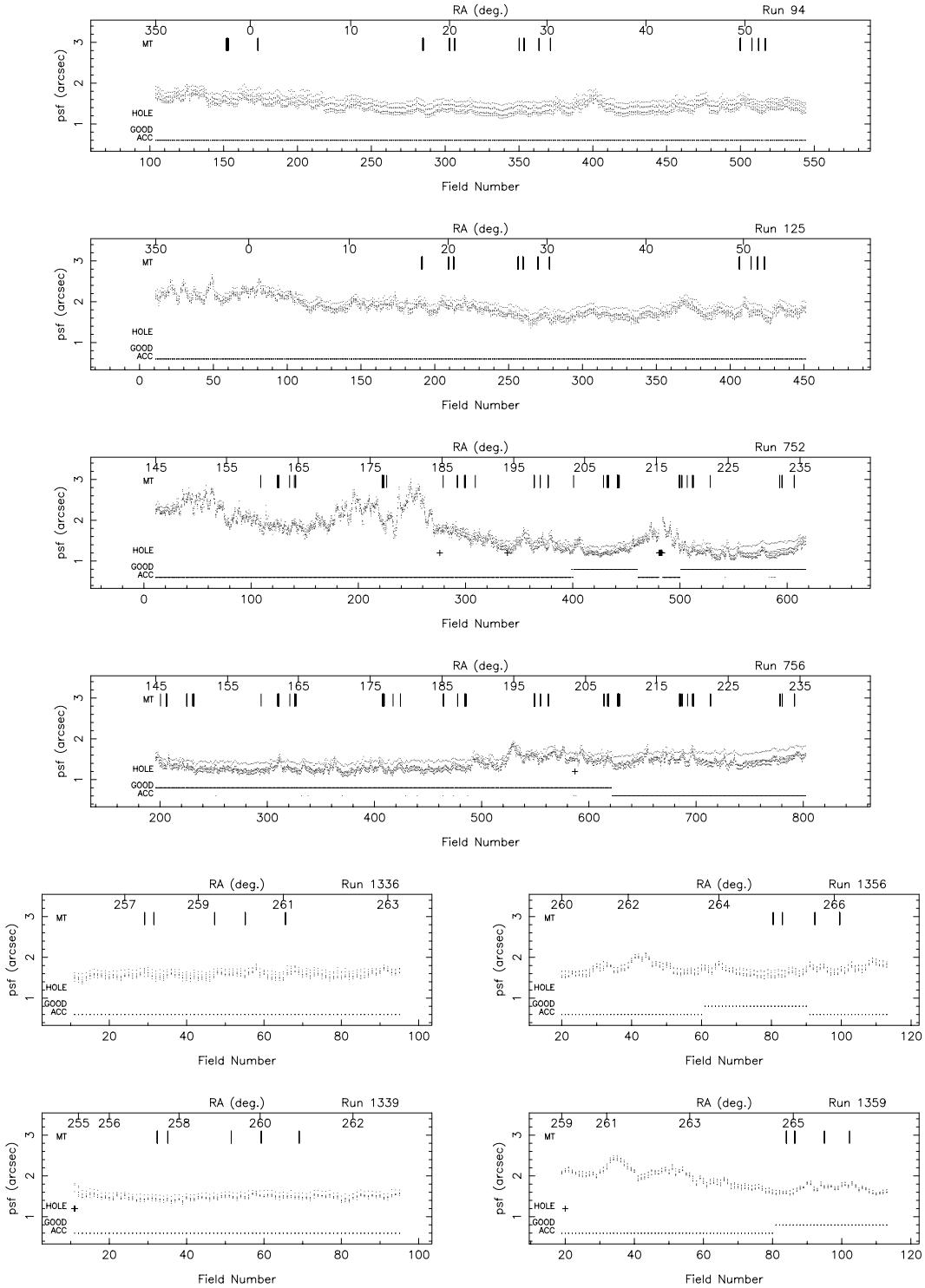


FIG. 8.—Seeing ( $\text{psfWidth}$ ) measured in the  $r$  filter in each of the imaging runs in the six camera columns, as a function of field number (lower axes) and  $\alpha$  (upper axes). A symbol for each field along the bottom indicates the quality of each field: Acceptable, Good, or Hole. Vertical marks indicate the location of MT secondary patches used for photometric calibration.

we assign it a new rerun number. The same rerun number in different runs does not necessarily refer to the same version of the pipelines.

**camCol** of class Segment (Table 15): This number, from 1 to 6, refers to the dewar, or column of photometric CCDs in the imaging camera, by which this object was imaged.

**fieldID** of class Field (Table 16): This refers to which field (§ 3.5) the object is in.

**objID** of class PhotoObj (Table 6): This is an identification number for the object that is unique within each field.

Several different versions of the pipelines were run on the imaging data for the EDR (Table 20). The rerun number for each run distinguishes a set of pipeline versions. Two rerun numbers are important for the EDR. The first rerun was used to select targets for spectroscopic observation. The sec-

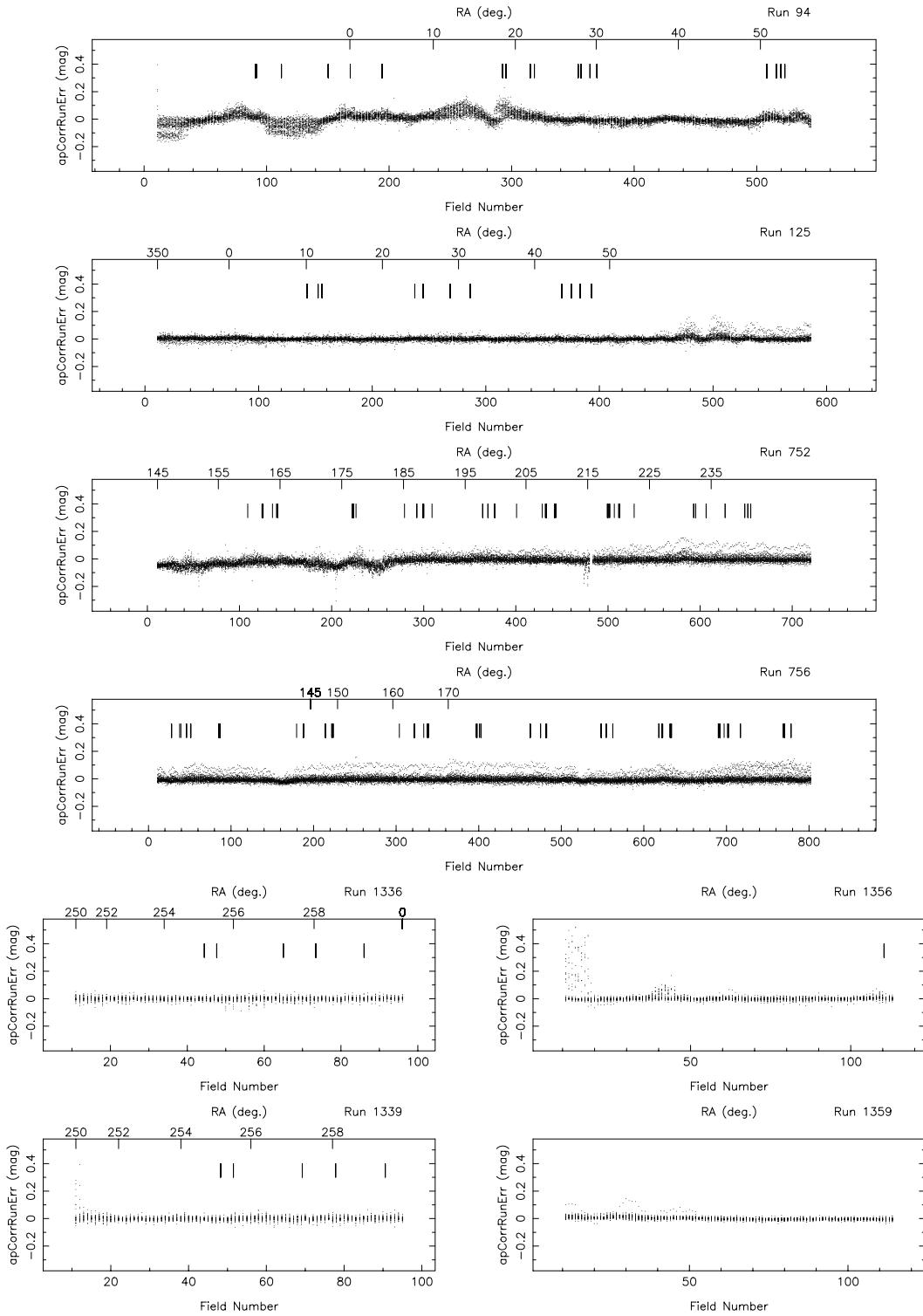


FIG. 9.—Aperture magnitude correction ( $\text{apCorrRunErr}$ ) for the five filters, in each of the imaging runs in the six camera columns, as a function of field number (*lower axes*) and  $\alpha$  (upper axes). Vertical marks indicate the location of MT secondary patches used for photometric calibration.

ond `rerun` used the most recent versions of the pipelines, and the results of this processing are distributed. Between reruns of `frames`, the `run`, `camCol`, and `fieldID` of detected objects do not change, because `frames` acts on individual frames. However, the `objID` does change. Future data releases will most likely be made with yet another `rerun` that uses the most current version of our pipelines and calibrations.

In addition, data-quality flags are set for each field, after we evaluate the processing from an entire run. These are discussed in § 4.6.

#### 4.4.2. Outline of frames

The pipeline analyzes the data one field (§ 3.5) at a time. Because information about each object is contained in five

separate frames, one for each filter, the five frames for each field are processed together.

Each frame has instrumental signatures (flat field, bias, cosmic rays, and bad columns) removed, and the global sky value from psp subtracted. The CCDs are known to be non-linear by of order 2% near saturation; this is *not* corrected for in the current version of the pipeline. Cosmic rays are found as objects with gradients between adjacent pixels substantially steeper than allowed by the PSF and are interpolated over. Note that our images are marginally well sampled in 1" seeing. Previously recognized bad columns are interpolated over using linear prediction (e.g., Rybicki & Press 1992), as are bleed trails from saturated stars.

Objects containing a (found and interpolated over) cosmic ray are flagged by having the mask bit CR set in flags for that band; objects with any interpolated pixels in them at all (due to bad columns or bleed trails) are flagged by having INTERP set. INTERP\_CENTER indicates that a pixel was interpolated over within 3 pixels of the center of the object, and PSF\_FLUX\_INTERP indicates that at least 20% of the PSF flux is interpolated; in rare cases, photometry of objects with these latter two flags set might be suspect.

Next, objects in the frame are detected and their properties measured in a four-step process in each band. First, an object finder is run to detect bright objects. In each band, the object finder detects pixels that are more than  $200\sigma$  (corresponding roughly to  $r^* = 17.5$ ) above the sky noise; only a single pixel need be over this threshold for an object to be detected at this stage. These objects are flagged as BRIGHT. The extended power-law wings of BRIGHT objects that are saturated are subtracted from the frame. Such stars are marked SUBTRACTED. Then the sky level is estimated by median-smoothing the frame image on a scale of approximately 100"; the resulting "local" sky image is subtracted from the frame (remember that a global sky determined on an entire frame has already been subtracted). This sky level is stored for each object for each band in the parameter sky (and an associated uncertainty skyErr), in units of asinh magnitudes (§ 4.4.5) per square arcsecond. For children of blends (§ 4.4.3), the sky parameter includes the contribution of its siblings.

Third, we find objects by smoothing the image with a Gaussian fit to the PSF and look for  $5\sigma$  peaks over the (smoothed) sky in each band. After objects are detected, they are "grown" more or less isotropically by an amount approximately equal to the radius of the seeing disk. We then define an object to be a connected set of pixels that are detected in at least one band. Note that all pixels in the object are subsequently used in the analysis in *every* band, whether or not they were originally detected in that band. The photometric pipeline never reports an upper limit for the detection of an object but, rather, carries out a proper measurement, with its error, for each of the varieties of flux listed in § 4.4.5.

Objects detected in a given band at this stage are flagged by setting the mask bit BINNED1 (Table 9) in flags of the PhotoObj class in that band. All pixel values in these BINNED1 objects are then replaced by the background level (with sky noise added in), the frame is rebinned into a  $2 \times 2$  image, and the object finder is run again. The resulting sample is flagged in a similar way with the BINNED2 mask, and pixel values in these objects are replaced with the background level. Finally, we rebin the original pixel data at  $4 \times 4$ , and objects found at this

stage are flagged BINNED4. The set of detected objects then consists of all objects with pixels flagged BINNED1, BINNED2, or BINNED4.

Fourth, the pipeline measures the properties of each object, including the position, as well as several measures of flux and shape, described more fully below. It attempts to determine whether each object actually consists of more than one object projected on the sky and, if so, to deblend such a "parent" object into its constituent "children," self-consistently across the bands (thus, all children have measurements in all bands). Then it again measures the properties of these individual children. Bright objects are measured twice: once with a global sky and no deblending run—this detection is flagged BRIGHT—and a second time with a local sky. For most purposes, only the latter is useful, and thus one should reject all objects flagged BRIGHT in compiling a sample of objects for study.

Other flag bits listed in Table 9 that are useful at this stage are

SATUR, indicating that at least one pixel in the object is flagged as saturated (the corresponding SATUR\_CENTER indicates that the saturated pixel or pixels are close to the center, to distinguish extended objects blended with saturated stars on their outskirts; unfortunately, this distinction is not as clean as we would like, because of improper handling of diffraction spikes); and

EDGE, indicating that the object overlaps the edge of the frame, which may affect the photometry of the object.

Finally, the pipeline outputs the measured quantities for each object, including all of the BRIGHT objects, all of the parents, and all of the children. In the following sections, we discuss how to interpret these outputs.

A typical frame at high Galactic latitudes contains of order 1000 objects, including of order five objects chosen deliberately in regions where no objects are detected; these are used to place sky fibers for spectroscopy and are classified as Sky. Repeat observations of a given area of sky in roughly 1.5" seeing shows that our 95% completeness limit for stars is  $u^* = 22.0$ ,  $g^* = 22.2$ ,  $r^* = 22.2$ ,  $i^* = 21.3$ , and  $z^* = 20.5$ ; the completeness drops to zero over roughly half a magnitude. These numbers are somewhat worse than quoted in the SDSS Project Book (see York et al. 2000), as that calculation assumed 1" seeing and slightly broader filters.

#### 4.4.3. Deblending of Objects

Once objects are detected, they are deblended by identifying individual peaks within each object, merging the list of peaks across bands, and adaptively determining the profile of images associated with each peak, which sum to form the original image in each band. The originally detected object is referred to as the "parent" object and has the flag BLENDED set if multiple peaks are detected; the final set of subimages of which the parent consists are referred to as the "children" and have the flag CHILD set. Note that all quantities in class PhotoObj (Table 6) are measured for *both* parent and child. For each child, parent gives the id of the parent (for parents themselves or isolated objects, this is set to the id of the BRIGHT counterpart if that exists; otherwise it is set to -1); for each parent, nchild gives the number of children an object has. Children are assigned the id numbers immediately after the id of the parent. Thus, if an object with id 23 is set as BLENDED and has nchild equal to

2, objects 24 and 25 will be set as CHILD and have parent equal to 23.

The list of peaks in the parent is trimmed to combine peaks (from different bands) that are too close to each other (if this happens, the flag PEAKS\_TOO\_CLOSE is set in the parent). If there are more than 25 peaks, only the most significant are kept, and the flag DEBLEND\_TOO\_MANY\_PEAKS is set in the parent.

In a number of situations, the deblender decides not to process a BLENDED object; in this case the object is flagged as NODEBLEND. Most objects with EDGE set are not deblended. The exceptions are when the object is large enough (larger than roughly an arcminute) that it will most likely not be completely included in the adjacent scan line either; in this case, DEBLENDED\_AT\_EDGE is set, and the deblender gives it its best shot. When an object is larger than half a frame, the deblender also gives up, and the object is flagged as TOO\_LARGE. Other intricacies of the deblending results are recorded in flags listed in Table 9.

On average, about 15%–20% of all detected objects are blended, and many of these are superpositions of galaxies that the deblender successfully treats by separating the images of the nearby objects. Thus, it is usually the childless (not BLENDED) objects that are of most interest for science applications. However, the versions of the pipelines used for the EDR occasionally deblend complex galaxies with large angular size, such as nearby face-on spiral galaxies, in a way the human eye would tend not to. Thus, some care is required in the analysis of samples of bright or large objects in the survey. Later versions of the deblender handle these cases much more gracefully, and future data releases will incorporate these improvements.

#### 4.4.4. Moving Objects

Main-belt asteroids have a proper motion of several arcseconds in the roughly 5 minutes it takes for an object to cross the imaging camera. This means that they will have a different centroid in the different photometric bands. If not taken into account, this could mean that they would be deblended into separate objects of unusual color, playing havoc with the target selection algorithms (§ 4.8). Thus, the deblender checks every object for consistency with uniform proper motion between the filters. In the PhotoObj class (Table 6), the quantities colv and rowv (and their associated errors) give the resulting proper motion (along the columns and rows of the CCDs, respectively) in units of degrees per day. The PSF photometry in each band is done on the object center, taking the motion into account, and, therefore, is properly measured (Ivezic et al. 2001).

Flag-bit values listed in Table 9 describe this processing. The MOVED flag indicates that the deblender considered whether to deblend the object as moving; it is not very useful. If the deblender actually deblended the object as moving, the flag DEBLENDED\_AS\_MOVING is set; otherwise the flag NODEBLEND\_MOVING is set. Note that an object can have a statistically significant motion without being deblended as such if the motion is small enough that the photometry would be fine without taking it into account. An object whose motion is not statistically significant is flagged STATIONARY, while an object whose motion is inconsistent with a straight line is flagged BAD\_MOVING\_FIT.

#### 4.4.5. Measurements of Flux

We have discussed how frames detects, deblends, and designates objects. This section and the next discuss the measurements applied to each resulting object. Each of the quantities described here has an associated estimated error measured as well, unless otherwise mentioned. In this subsection, we discuss the various measurements made of the flux in each object.

##### 4.4.5.1. The asinh Magnitude Scale

We begin by describing the magnitude scale that the SDSS uses. Unless otherwise specified (the most important exceptions being petroMag and modelMag, to obtain self-consistent colors), the measures discussed here are applied independently in each bandpass. Magnitudes within the SDSS are expressed as inverse hyperbolic sine (or “asinh”) magnitudes, described in detail by Lupton, Gunn, & Szalay (1999). The transformation from linear flux measurements to asinh magnitudes is designed to be virtually identical to the standard astronomical magnitude (Pogson 1856) at high signal-to-noise ratio but to behave reasonably at low signal-to-noise ratio and even at negative values of flux, where the logarithm in the Pogson magnitude fails. This allows us to measure a flux even in the absence of a formal detection; we quote no upper limits in our photometry.

The asinh magnitudes are characterized by a softening parameter  $b$ , the typical  $1\sigma$  noise of the sky in a PSF aperture in 1" seeing. The relation between detected flux  $f$  and asinh magnitude  $m$  is

$$m = -\frac{2.5}{\ln 10} \left[ \operatorname{asinh} \left( \frac{f/f_0}{2b} \right) + \ln b \right] \quad (6)$$

(see eq. [3] of Lupton et al. 1999), where  $f_0$  sets the zero point of the magnitude scale (see the discussion in § 4.5, and especially § 4.5.4). The quantity  $b$  is measured relative to  $f_0$  and is thus dimensionless; it is given in Table 21, along with the asinh magnitude associated with a zero-flux object. The table also lists the flux corresponding to  $10f_0b$ , above which the asinh magnitude and the traditional logarithmic magnitude differ by less than 1% in flux.

TABLE 21  
asinh MAGNITUDE SOFTENING PARAMETERS

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

NOTE.—These values of the softening parameter  $b$  are set to be approximately  $1\sigma$  of the sky noise; thus, only low signal-to-noise ratio 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. (Note that the definition  $b$  has been modified since its mention in the appendix of Fan et al. 1999. There,  $b$  had units of  $\text{DN s}^{-1}$ ; now it is a dimensionless quantity.)

#### 4.4.5.2. PSF Magnitudes

For isolated stars, which are well described by the PSF, the optimal measure of the total flux is determined by fitting a PSF model to the object. In practice, we do this by sinc-shifting the image of a star so that it is exactly centered on a pixel and then fitting a Gaussian model of the PSF to it. This fit is carried out on the local PSF K-L model (§ 4.3) at each position as well; the difference between the two is then a local aperture correction, which gives a corrected PSF magnitude. Finally, we use bright stars to determine a further aperture correction to a radius of 7''.4 as a function of seeing and apply this to each frame for its seeing. This involved procedure is necessary to take into account the full variation of the PSF (measured in the postage-stamp pipeline, described above) across the field, including the low signal-to-noise ratio wings. Empirically, this reduces the seeing dependence of the photometry to below 0.02 mag for seeing as poor as 2''. The resulting magnitude is stored in the quantity `psfMag`. As mentioned above, the flag `PSF_FLUX_INTERP` warns that the PSF photometry might be suspect. The flag `BAD_COUNTS_ERROR` warns that because of interpolated pixels, the error may be underestimated.

The PSF errors include contributions from photon statistics and uncertainties in the PSF model and aperture correction, although they do not include uncertainties in photometric calibration (§ 4.5). Repeat observations show that these errors are probably underestimated by 10%–20%.

#### 4.4.5.3. Fiber Magnitudes

The flux contained within the aperture of a spectroscopic fiber (3'' in diameter) is calculated in each band and stored in `fiberMag`. Note that no correction for seeing is applied to this measure of the magnitude. For children of deblended galaxies, some of the pixels within a 1''.5 radius may belong to other children. In this case, the fiber magnitudes can be rather misleading, as they will not reflect the amount of light that the spectrograph will see. For future data releases, we will calculate the true flux within a fiber diameter, including all light from the parent that falls in the aperture centered at the location of the child. We will also correct the detected flux to a fiducial value of the seeing.

#### 4.4.5.4. Petrosian Magnitudes

For galaxy photometry, measuring flux is more difficult than for stars, because galaxies do not all have the same radial surface brightness profile, and they have no sharp edges. In order to avoid biases, we wish to measure a constant fraction of the total light, independent of the position and distance of the object. To satisfy these requirements, the SDSS has adopted a modified form of the Petrosian (1976) system, measuring galaxy fluxes within a circular aperture whose radius is defined by the shape of the azimuthally averaged light profile.

We define the “Petrosian ratio”  $\mathcal{R}_P$  at a radius  $r$  from the center of an object to be the ratio of the local surface brightness in an annulus at  $r$  to the mean surface brightness within  $r$ , as described by Blanton et al. (2001) and Yasuda et al. (2001):

$$\mathcal{R}_P(r) \equiv \frac{\int_{0.8r}^{1.25r} dr' 2\pi r' I(r') / [\pi(1.25^2 - 0.8^2)r^2]}{\int_0^r dr' 2\pi r' I(r') / (\pi r^2)}, \quad (7)$$

where  $I(r)$  is the azimuthally averaged surface brightness profile.

The Petrosian radius  $r_P$  is defined as the radius at which  $\mathcal{R}_P(r_P)$  equals some specified value  $\mathcal{R}_{P,\text{lim}}$ , set to 0.2 in our case. The Petrosian flux in any band is then defined as the flux within a certain number  $N_P$  (equal to 2.0 in our case) of Petrosian radii:

$$F_P \equiv \int_0^{N_P r_P} 2\pi r' dr' I(r'). \quad (8)$$

In the SDSS five-band photometry, the aperture in all bands is set by the profile of the galaxy in the  $r$  band alone. This procedure ensures that the color measured by comparing  $F_P$  in different bands is measured through a consistent aperture.

The aperture  $2r_P$  is large enough to contain nearly all of the flux for typical galaxy profiles, but small enough that the sky noise in  $F_P$  is small. Thus, even substantial errors in  $r_P$  cause only small errors in the Petrosian flux (typical statistical errors near the spectroscopic flux limit of  $r^* \approx 17.7$  are less than 5%), although these errors are correlated.

The Petrosian radius in each band is the parameter `petroRad`, and the Petrosian magnitude in each band (calculated, remember, using only `petroRad` for the  $r$  band) is the parameter `petroMag`.

In practice, there are a number of complications associated with this definition, because noise, substructure, and the finite size of objects can cause objects to have no Petrosian radius, or more than one. Those with more than one are flagged as `MANYPETRO`; the largest one is used. Those with none have `NOPETRO` set. Most commonly, these objects are faint ( $r^* > 20.5$  or so), and the Petrosian ratio becomes unmeasurable before dropping to the limiting value of 0.2; these have `PETROFAINT` set and have their “Petrosian radii” set to the default value of the larger of 3'' and the outermost measured point in the radial profile. Finally, a galaxy with a bright stellar nucleus, such as a Seyfert galaxy, can have a Petrosian radius set by the nucleus alone; in this case, the Petrosian flux misses most of the extended light of the object. This happens quite rarely, but one dramatic example in the EDR data is the Seyfert galaxy NGC 7603 (=Arp 92), at  $\alpha = 23^{\text{h}}18^{\text{m}}56\overset{\text{s}}{.}6$ ,  $\delta = +00^{\circ}14'38''$  (J2000).

How well does the Petrosian magnitude perform as a reliable and complete measure of galaxy flux? Theoretically, the Petrosian magnitudes defined here should recover essentially all of the flux of an exponential galaxy profile and about 80% of the flux for a de Vaucouleurs profile. As shown by Blanton et al. (2001), this fraction is fairly constant with axis ratio, while as galaxies become smaller (as a result of worse seeing or greater distance), the fraction of light recovered becomes closer to that fraction measured for a typical PSF, about 95% in the case of the SDSS. This implies that the fraction of flux measured for exponential profiles decreases while the fraction of flux measured for de Vaucouleurs profiles increases as a function of distance. However, for galaxies in the spectroscopic sample ( $r^* < 17.7$ ), these effects are small; the Petrosian radius measured by frames is extraordinarily constant in physical size as a function of redshift.

#### 4.4.5.5. Model Fits

Just as the PSF magnitudes are optimal measures of the fluxes of stars, the optimal measure of the flux of a galaxy

would use a matched galaxy model. With this in mind, the code fits two models to the two-dimensional image of each object in each band: a pure de Vaucouleurs profile,

$$I(r) = I_0 \exp \{-7.67[(r/r_{\text{eff}})^{1/4}]\} \quad (9)$$

(truncated beyond  $7r_{\text{eff}}$  to smoothly go to zero at  $8r_{\text{eff}}$ , and with some softening within  $r = r_{\text{eff}}/50$ ), and a pure exponential profile,

$$I(r) = I_0 \exp (-1.68r/r_{\text{eff}}) \quad (10)$$

(truncated beyond  $3r_{\text{eff}}$  to smoothly go to zero at  $4r_{\text{eff}}$ ), each of arbitrary axis ratio and position angle. Although for large objects it is possible and even desirable to fit more complicated models (e.g., bulge plus disk), the computational expense to compute them is not justified for the majority of the detected objects.<sup>68</sup> The models are convolved with a double-Gaussian fit to the PSF, which is provided by psp. Residuals between the double-Gaussian and the full K-L PSF model are added on for just the central PSF component of the image.

At bright magnitudes ( $r^* < 18$ ), the model magnitudes are a poor measure of the total flux of the galaxy, because of the fact that the fits are restricted to the central parts of objects (Strateva et al. 2001). This issue will be addressed in future data releases.

These fitting procedures yield the quantities  $r_{\text{deV}}$  and  $r_{\text{exp}}$ , the effective radii of the models;  $ab_{\text{deV}}$  and  $ab_{\text{exp}}$ , the axis ratio of the best-fit models; and  $\text{phi}_{\text{deV}}$  and  $\text{exp\_phi}$ , the position angles of the ellipticity (in degrees east of north). Note that these quantities correctly model the effects of the PSF. Also recorded are the likelihoods associated with each model from the  $\chi^2$  fit,  $\text{deV\_L}$  and  $\text{exp\_L}$ , and the total magnitudes associated with each fit,  $\text{deVMag}$  and  $\text{expMag}$ . Errors for each of these quantities (which are based only on photon statistics) are also reported. We apply aperture corrections to make these model magnitudes equal the PSF magnitudes in the case of an unresolved object.

In order to measure unbiased colors of galaxies, we measure their flux through equivalent apertures in all bands. We choose the model (exponential or de Vaucouleurs) of higher likelihood in the  $r$  filter, and we apply that model (i.e., allowing only the amplitude to vary) in the other bands after convolving with the appropriate PSF in each band. The resulting magnitudes are termed `modelMag`. The resulting estimate of galaxy color will be unbiased in the absence of color gradients. Systematic differences from Petrosian colors are in fact often seen as a result of color gradients, in which case the concept of a global galaxy color is somewhat ambiguous. For faint galaxies, the model colors have appreciably higher signal-to-noise ratio than do the Petrosian colors.

Because of the way in which model fits are carried out, there is some weak discretization of model parameters, especially  $r_{\text{exp}}$  and  $r_{\text{deV}}$ . Also note that very occasionally the quoted axis ratios are given negative values. These two problems will be addressed in subsequent data releases.

<sup>68</sup> In order to keep up with the data processing, the frames pipeline uses only 15–20 ms per object per band on a 1 GHz Pentium III machine.

TABLE 22  
CONVERSION FROM  
 $E(B-V)$  TO EXTINCTION

Band	$A_{\text{filter}}/E(B-V)$
$u.....$	5.155
$g.....$	3.793
$r.....$	2.751
$i.....$	2.086
$z.....$	1.479

#### 4.4.5.6. Reddening

Reddening corrections in magnitudes at the position of each object, `reddening`, are computed following Schlegel, Finkbeiner, & Davis (1998). These corrections are *not* applied to the magnitudes in the databases. Conversions from  $E(B-V)$  to total extinction  $A_\lambda$ , assuming a  $z=0$  elliptical galaxy spectral energy distribution, are tabulated in Table 22.

#### 4.4.5.7. Which Magnitudes Should One Use?

Faced with this array of different magnitude measurements to choose from, which one is appropriate in which circumstances? We cannot give any guarantees of what is appropriate for the science *you* want to do, but here we present some examples, where we use the general guideline that one usually wants to maximize some combination of signal-to-noise ratio, fraction of the total flux included, and freedom from systematic variations with observing conditions and distance:

*Photometry of bright stars.*—If the objects are bright enough, add up all of the flux from the profile `profMean` and generate a large-aperture magnitude. We recommend using the first seven annuli.

*Photometry of distant quasars.*—These will be unresolved and therefore have images consistent with the PSF. For this reason, `psfMag` is unbiased and optimal.

*Colors of stars.*—Again, these objects are unresolved, and `psfMag` is the optimal measure of their brightness.

*Photometry of nearby galaxies.*—Galaxies bright enough to be included in our spectroscopic sample have relatively high signal-to-noise ratio measurements of their Petrosian magnitudes. Since these magnitudes are model independent and yield a large fraction of the total flux, roughly constant with redshift, `petroMag` is the measurement of choice for such objects. In fact, the noise properties of Petrosian magnitudes remain good to  $r^* = 20$  or so.

*Photometry of distant galaxies.*—For the faintest galaxies, estimates of the Petrosian magnitudes become very noisy. Under these conditions, the `modelMag` is usually a more reliable estimate of the galaxy flux. In addition, these magnitudes account for the effects of local seeing and thus are less dependent on local seeing variations. This property also is desirable for the faintest, smallest objects. The model colors for galaxies are also unbiased, as mentioned above.

Of course, it would *not* be appropriate to study the evolution of galaxies and their colors by using Petrosian magnitudes for bright galaxies and model magnitudes for faint galaxies.

Finally, we note that azimuthally averaged radial profiles are provided, as described below, and can easily be used to

create circular-aperture magnitudes of any desired type. For instance, to study a large dynamic range of galaxy fluxes, one could measure alternative Petrosian magnitudes with parameters tuned such that the Petrosian flux includes a small fraction of the total flux but yields higher signal-to-noise ratio measurements at faint magnitudes.

#### 4.4.6. Measurements of Shape and Morphology

The frames pipeline also provides several characterizations of the shape and morphology of an object.

##### 4.4.6.1. frames's "Type" Determination

The frames pipeline provides a simple star-galaxy separator in its type parameters (provided separately for each band) and its `objc_type` parameters (one value per object); these are set to 3 (galaxy) or 6 (star). In particular, the following simple cut works at the 95% confidence level for our data to  $r^* = 21$  and even somewhat fainter:

$$\text{psfMag} - (\text{deV\_L} > \text{exp\_L}) ? \text{deVMag} : \text{expMag} > 0.145. \quad (11)$$

If satisfied, `type` is set to GALAXY for that band; otherwise, `type` is set to STAR. The global type `objc_type` is set according to the same criterion, applied to the summed fluxes from all bands in which the object is detected.

Experimentation has shown that simple variants on this scheme, such as defining galaxies as those objects classified as such in any two of the three high signal-to-noise ratio bands (namely,  $g$ ,  $r$ , and  $i$ ), work better in some circumstances. This scheme occasionally fails to distinguish pairs of stars with separation small enough ( $<2''$ ) that the deblender does not split them; it also occasionally classifies Seyfert galaxies with particularly bright nuclei as stars.

Additional information to refine the star-galaxy separation further may be used, depending on scientific application. For example, Scranton et al. (2001) advocate applying a Bayesian prior to the above difference between the PSF and exponential magnitudes, depending on seeing and using prior knowledge about the counts of galaxies and stars with magnitude.

##### 4.4.6.2. Radial Profile

The frames pipeline extracts an azimuthally averaged radial surface brightness profile. In the catalogs, it is given as the average surface brightness in a series of annuli (whose outer radii are listed in Table 7). This quantity is in units of "maggies" per square arcsecond, where a maggie is a linear measure of flux; 1 maggie has an AB magnitude of 0 (thus, a surface brightness of 20 mag arcsec $^{-2}$  corresponds to  $10^{-8}$  maggies arcsec $^{-2}$ ). In the class `Profile` (Table 6), the number of annuli for which there is a measurable signal is listed as `nprof`, the mean surface brightness is listed as `profMean`, and the error is listed as `profErr`. This error includes both photon noise and the small-scale "bumpiness" in the counts as a function of azimuthal angle.

When converting the `profMean` values to a local surface brightness, it is *not* the best approach to assign the mean surface brightness to some radius within the annulus and then linearly interpolate between radial bins. Do not use smoothing splines, as they will not go through the points in the cumulative profile and thus (obviously) will not conserve flux. What frames does, for example, in determining the

Petrosian ratio, is to fit a taut spline to the *cumulative* profile and then differentiate that spline fit, after transforming both the radii and cumulative profiles with asinh functions. We recommend doing the same here.

##### 4.4.6.3. Surface Brightness and Inverse Concentration Index

The frames pipeline also reports the radii containing 50% and 90% of the Petrosian flux for each band, `petroR50` and `petroR90`, respectively. The usual characterization of surface brightness in the target selection pipeline of the SDSS is the mean surface brightness within `petroR50`.

It turns out that the ratio of `petroR50` to `petroR90`, the so-called inverse concentration index, is correlated with morphology (Shimasaku et al. 2001; Strateva et al. 2001). Galaxies with a de Vaucouleurs profile have an inverse concentration index of around 0.3; exponential galaxies have an inverse concentration index of around 0.43. Thus, this parameter can be used as a simple morphological classifier.

An important caveat when using these quantities is that they are *not* corrected for seeing. This causes the surface brightness to be underestimated, and the inverse concentration index to be overestimated, for objects of size comparable to the PSF. The amplitudes of these effects, however, are not yet well characterized.

##### 4.4.6.4. Model Fit Likelihoods and Parameters

In addition to the model and PSF magnitudes described above, the likelihoods `deV_L`, `exp_L`, and `star_L` are also calculated by frames. These are the probabilities of achieving the measured  $\chi^2$  for the de Vaucouleurs, exponential, and PSF fits, respectively. If one wishes to make use of this trinary scheme to classify objects, calculation of the fractional likelihoods is recommended:

$$f(\text{deV\_L}) = \frac{\text{deV\_L}}{\text{deV\_L} + \text{exp\_L} + \text{star\_L}}, \quad (12)$$

and similarly for  $f(\text{exp\_L})$  and  $f(\text{star\_L})$ . A fractional likelihood greater than 0.5 for any of these three profiles is generally a good threshold for object classification. This works well in the range  $18 < r^* < 21.5$ ; at the bright end, the likelihoods have a tendency to underflow to zero, which makes them less useful. In particular, `star_L` is often zero for bright stars. For future data releases, we will incorporate improvements to the model fits to give more meaningful results at the bright end.

##### 4.4.6.5. Ellipticity

The model fits yield an estimate of the axis ratio and position angle of each object, but it is useful to have model-independent measures of ellipticity. In the data released here, frames provides two further measures of ellipticity, one based on second moments, the other based on the ellipticity of a particular isophote. The model fits do correctly account for the effect of the seeing, while the methods presented here do not.

The first method measures flux-weighted second moments, defined as

$$M_{xx} \equiv \left\langle \frac{x^2}{r^2} \right\rangle, \quad M_{yy} \equiv \left\langle \frac{y^2}{r^2} \right\rangle, \quad M_{xy} \equiv \left\langle \frac{xy}{r^2} \right\rangle. \quad (13)$$

In the case that the object's isophotes are self-similar ellip-

ses, one can show that

$$Q \equiv M_{xx} - M_{yy} = \frac{a-b}{a+b} \cos 2\phi, \quad U \equiv M_{xy} = \frac{a-b}{a+b} \sin 2\phi, \quad (14)$$

where  $a$  and  $b$  are the semimajor and semiminor axes and  $\phi$  is the position angle.  $Q$  and  $U$  are Q and U in the class PhotoObj and are referred to as “Stokes parameters.” They can be used to reconstruct the axis ratio and position angle, measured relative to row and column of the CCDs. This is equivalent to the normal definition of position angle (east of north), for the scans on the equator. The performance of the Stokes parameters are not ideal at low signal-to-noise ratio. For future data releases, frames will also output variants of the adaptive shape measures used in the weak-lensing analysis of Fischer et al. (2000), which are closer to optimal measures of shape for small objects.

A second measure of ellipticity is given by measuring the ellipticity of the 25 mag arcsec $^{-2}$  isophote (in all bands). In detail, frames measures the radius of a particular isophote as a function of angle and Fourier-expands this function. It then extracts from the coefficients the centroid, major and minor axes, position angle, and average radius of the isophote in question.<sup>69</sup> It also reports the derivative of each of these quantities with respect to isophote level, necessary to recompute these quantities if the photometric calibration changes.<sup>70</sup>

#### 4.4.7. Known Problems in frames

Throughout this section, we have listed various known problems with the processing, many of which will be fixed in future releases of the data. We now list problems that were not otherwise mentioned earlier.

1. There are several quantities that have not yet been finalized and are currently placeholders in the pipeline outputs. These include error estimates for PetroR50 and PetroR90, all errors associated with isophotal quantities, a texture parameter to measure the small-scale roughness of an object, and fracPSF, the fraction of light included in the PSF. Quantities that are not calculated are designated  $-9999$  in the database, while *errors* that are not calculated are designated  $-1000$ . Note that some quantities are not calculated, for example, for objects very close to the edge of a frame.

2. We have not done a thorough testing of every one of the outputs of the pipeline. Quantities such as psfMag and petroMag, which are used in target selection (§ 4.8), are very extensively tested, while quantities such as the parameters of the elliptical isophote fit, or the errors in the model fit parameters, have not been as carefully vetted.

3. Because of difficulties in cosmic-ray rejection, one should be suspicious of objects detected in only one band. This is particularly problematic in  $z$ , where the occasional cosmic ray can splatter in the thick chip to be indistinguishable from the PSF (see the discussion in Fan et al. 2001a).

<sup>69</sup> Placeholders exist in the database for the errors on each of these quantities, but they are not currently calculated.

<sup>70</sup> Note that these isophotal quantities are the only ones that require a priori photometric calibrations before frames is run. They use a preliminary calibration, which is only accurate to 0.05 mag or so.

Future releases of frames will set a flag MAYBE\_CR to flag ambiguous cases.

4. Electronics cross talk can produce faint ghosts of saturated stars in  $z$ . Draw a line down the central column of the CCD; the ghost will appear at the mirror position relative to this line. These will be flagged in future releases.

5. The code does not subtract diffraction spikes from stars (although bleed trails are interpolated over). These will be subtracted in future releases.

6. There are a variety of faint ghosts due to reflection off the edges of the filters. Thus, one should be suspicious of the reality of very low surface brightness (fainter than 24 mag arcsec $^{-2}$  in  $r$ ) extended features close to the edges of frames, especially in the  $u$  band.

7. Similarly, the deblender occasionally pulls off low surface brightness features from large galaxies. This problem is much reduced with improvements to the deblender in subsequent versions of the pipeline.

8. The proper-motion errors are occasionally underestimated (yielding a few quasars with nominally significant proper motions!). This is due to uncertainties in the astrometric offsets between bands and will be fixed in future releases. However, overlapping scans have shown that the proper-motion errors for the majority of asteroids are *overestimated* by roughly a factor of 2.

9. Ghosting in the  $u$  filter causes systematic errors in the flat fields in some of the runs, affecting the photometry by up to 10%. This is discussed in detail in § 4.5.3 below.

Finally, a disclaimer: we have tested the outputs of the pipeline to the best of our ability, but there are no doubt subtle problems not listed here that we are currently unaware of. For example, many of the quoted error estimates have not yet been tested adequately. We are eager for feedback from the community if any problems are found.

#### 4.5. Photometric Calibration

The objective of the photometric calibration process is to tie the SDSS imaging data to an AB magnitude system, and specifically to the “natural system” of the 2.5 m telescope, defined by the photon-weighted effective wavelengths of each combination of SDSS filter, CCD response, telescope transmission, and atmospheric transmission at a reference air mass of 1.3 as measured at APO.

The calibration process ultimately involves combining data from three telescopes: the USNO 40 inch, on which our primary standards were first measured (see below), the SDSS PT (§ 3.4), and the SDSS 2.5 m telescope. At the beginning of the survey, it was expected that there would be a single  $u'g'r'i'z'$  system. However, in the course of processing the SDSS data, the unpleasant discovery was made that the filters in the 2.5 m telescope have significantly different effective wavelengths than the filters in the PT and at USNO. These differences have been traced to the fact that the short-pass interference films on the 2.5 m camera live in the same vacuum as the detectors, and the resulting dehydration of the films decreases their effective refractive index. This results in blueward shifts of the red edges of the filters by about 2.5% of the cutoff wavelength, and consequent shifts of the *effective* wavelengths of order half that. The USNO filters are in ambient air, and the hydration of the films exhibits small temperature shifts; the PT filters are kept in stable, very dry air and are in a condition about halfway between ambient and the very stable vacuum state. The

rather subtle differences between these systems are describable by simple linear transformations with small color terms for stars of not-too-extreme color but, of course, cannot be so transformed for very cool objects or objects with complex spectra. Since standardization is done with stars, this is not a fundamental problem, once the transformations are well understood.

It is these subtle issues that give rise to our somewhat awkward nomenclature for the PT and 2.5 m filter systems. The calibration is provisional because the zero points are not quite on an AB system yet. For many purposes, the difference between the USNO and the 2.5 m systems are of no consequence. We believe that the photometric calibration of the EDR is uniform to of order 3% in  $g^*$ ,  $r^*$ , and  $i^*$ , and 5% in  $u^*$  and  $z^*$  (all numbers are rms), while the absolute calibration in janskys is uncertain by as much as 10%.

In this section, we describe the process by which photometric calibration is done (§ 4.5.1), requiring analysis of data from the PT (§ 4.5.2), which is then combined with the 2.5 m imaging data (§ 4.5.3). The tests we have carried out to give us confidence in the above numbers, and caveats in the calibration, are described in § 4.5.4. We do not aim here to give a full pedagogical description of the difficulties of absolute photometry in the SDSS system; that will be presented in future papers.

#### 4.5.1. Overview of the Photometric Calibration in the SDSS

The photometric calibration of the SDSS imaging data is a multistep process, as a result of the fact that the images from the 2.5 m telescope saturate at  $r^* \approx 14$ , fainter than typical spectrophotometric standards, combined with the fact that observing efficiency would be greatly impacted if the 2.5 m needed to interrupt its routine scanning in order to observe separate calibration fields.

The first step involved setting up a primary standard-star network of 163 stars distributed around the northern sky (Smith et al. 2002). These stars were selected from a variety of sources and span a range in color, air mass, and right ascension. They were observed repeatedly over a period of 2 years using the USNO 40 inch telescope located in Flagstaff, Arizona. These observations are tied to an absolute flux system by the single F0 subdwarf star BD +17°4708, whose absolute fluxes in the SDSS filters are taken from Fukugita et al. (1996). As noted above, the photometric system defined by these stars is called the  $u'g'r'i'z'$  system. A table containing the calibrated magnitudes for these standard stars is available from our Web sites.

Most of these primary standards have brightnesses in the range  $r^* \approx 8\text{--}13$  and would saturate the 2.5 m telescope's imaging camera in normal operations. Therefore, a set of 1520  $41.5' \times 41.5'$  transfer fields, called *secondary patches*, have been positioned throughout the survey area. These secondary patches are observed with the PT (§ 3.4); their size is set by the field of view of the PT camera. These secondary patches are grouped into sets of four. Each set spans the full set of 12 scan lines of a survey stripe along the width of the stripe, and the sets are spaced along the length of a stripe at roughly  $15^\circ$  intervals. The locations of the patches used for calibrating the imaging runs for the EDR are indicated in Figure 8. The patches are observed by the PT in parallel with observations of the primary standards and processed using the monitor telescope pipeline (§ 4.5.2); the calibrated patches are then used to calibrate the 2.5 m's imaging data

via the final calibrations pipeline (§ 4.5.3) to the 2.5 m preliminary  $u^*g^*r^*i^*z^*$  system.

#### 4.5.2. Monitor Telescope Pipeline

The PT has two main functions: it measures the atmospheric extinction on each clear night based on observations of primary standards at a variety of air masses, and it calibrates secondary patches in order to determine the photometric zero point of the 2.5 m imaging scans. The extinction must be measured on each night the 2.5 m is scanning, but the corresponding secondary patches can be observed on any photometric night and need not be coincident with the image scans that they will calibrate.

The monitor telescope pipeline (`mtpipe`), so called for historical reasons, processes the PT data. It performs three basic functions:

1. It bias-subtracts and flat-fields the images, and performs aperture photometry.
2. It identifies primary standards in the primary standard-star fields and computes a transformation from the aperture photometry to the primary standard-star system.
3. It applies the photometric solution to the stars in the secondary patch fields.

The transformation equation between the instrumental aperture photometry  $u'_{\text{inst}}$  and the reduced magnitude  $u'$  is given by

$$u'_{\text{inst}} = u' + a_u + b_u(u' - g') + k_u X + c_u[(u' - g') - (u' - g')_{\text{ref}}](X - X_{\text{ref}}), \quad (15)$$

and similarly in  $g'$ ,  $r'$ ,  $i'$ , and  $z'$  (both  $i'$  and  $z'$  use  $i' - z'$  as the fiducial color in the equations). Here  $a$  is the zero point,  $k$  is the primary extinction coefficient,  $X$  is the air mass,  $b$  is the first-order color term coefficient,  $c$  is the second-order extinction term coefficient,  $X_{\text{ref}} = 1.3$  is a reference air mass, and  $(u' - g')_{\text{ref}}$  is a reference color. The values for these last two parameters are chosen to keep the “ $c$ ” term small and are defined to be the estimated average air mass of the survey,  $\langle X \rangle = 1.3$ , and the “cosmic color,” respectively, as listed in Table 23. The cosmic color values are the mean colors of 4428 objects with  $19 < r^* < 20$  in survey run 752, camera column 3, fields 11–100. This area is on the celestial equator at a Galactic latitude of about  $42^\circ$ .

For the calibrations presented in the EDR, the values for the instrumental (“ $b$ ”) coefficients have been set to zero, and the values for the second-order extinction (“ $c$ ”) coefficients have been set to the global values derived in the initial setup of the  $u'g'r'i'z'$  primary standard-star network (Table 24; Smith et al. 2002). For each filter, a single value for the photometric zero point (“ $a$ ”) is solved for each night, but typically a separate  $k$  is determined for each 3 hr block throughout the night. Typical values are given in Table 24.

TABLE 23  
THE COSMIC COLORS

Color	Value	Scatter
$u^* - g^*$ .....	1.42	$\pm 0.82$
$g^* - r^*$ .....	1.11	$\pm 0.46$
$r^* - i^*$ .....	0.48	$\pm 0.34$
$i^* - z^*$ .....	0.35	$\pm 0.22$

TABLE 24  
EXTINCTION TERMS AND SECOND-ORDER  
EXTINCTION COEFFICIENTS

Filter	<i>k</i>	<i>c</i>
<i>u</i> .....	$0.55 \pm 0.10$	$-0.021 \pm 0.003$
<i>g</i> .....	$0.15 \pm 0.05$	$-0.016 \pm 0.003$
<i>r</i> .....	$0.10 \pm 0.04$	$-0.004 \pm 0.003$
<i>i</i> .....	$0.08 \pm 0.02$	$+0.006 \pm 0.003$
<i>z</i> .....	$0.06 \pm 0.02$	$+0.003 \pm 0.003$

The net effect of forcing the *b*-terms to zero is that the secondary patches are calibrated on the natural system of the PT, except that the zero points in each filter are slightly offset from a true AB system. The offset for the *u'* band is given approximately by  $b\langle u' - g' \rangle$ , where the quantity in brackets is the mean color of all standard stars observed on a given night; offsets for other bands are given by similar expressions. These offsets amount to at most a couple of percent and are not applied to the EDR data.

#### 4.5.3. Final Calibration Pipeline

The final calibration pipeline works much like `mtpipe`, computing the transformation between aperture photometry (or other photometry) as observed by the 2.5 m telescope and the final SDSS photometric system. The pipeline matches stars between a camera column of 2.5 m data and an overlapping secondary patch. Each camera column of 2.5 m data is calibrated individually. There are of order 100 stars in each patch in the appropriate color and magnitude range in the overlap. The locations of these secondary patches are indicated in Figure 8.

The transformation equations have the same form as those used by `mtpipe`, i.e., equation (15). Once again the *b*-coefficient is set to zero. The extinction coefficient is taken from PT observations on the same night, linearly interpolated in time when multiple extinction determinations are available. Second-order extinction corrections are included for completeness but, in practice, do not affect the photometry by more than 0.01 mag. A single zero point *a* is computed from stars on all patches that overlap a given CCD in a given run. Observations are weighted by their estimated errors, and  $\sigma$ -clipping is used to reject outliers. At one time it was thought that a time-dependent zero point might be needed to account for the fact that the 2.5 m camera and corrector lenses rotate relative to the telescope mirrors and optical structure; however, it now appears that any variations in throughput are small compared with inherent fluctuations in the calibration of the patches themselves. The statistical error in the zero point is constrained to be less than 1.35% in  $u^*$  and  $z^*$  and 0.9% in  $g^*r^*i^*$ . However, systematic errors can be larger; we describe some of the tests of the consistency of our photometry in § 4.5.4.

One of the diagnostics of systematic problems is to plot residuals in the individual star observations as a function of magnitude or CCD column. Figure 10 shows an example, which indicates a problem with ghosting internal to the 2.5 m telescope and system. This is most apparent in the *u* band and is due to ghosting between the CCD and the *u* filter combined with vignetting of the out-of-focus ghost image. At the edge of the field of view, especially columns 1 and 6 of the 2.5 m imaging camera, this effect gives rise to a

CCD column-dependent magnitude error of up to 10% in  $u^*$  for a small range of CCD columns in the 2.5 m imaging camera. The effect is present to a smaller degree (no more than 3%) in the other filters. The effect exists in the EDR in runs 752, 756, 1336, 1339, 1356, and 1359, and it will be addressed for future data releases.

For the first two runs included in the EDR (94 and 125), the full telescope baffling system was not in place. Stray light entering the telescope affected the median-sky flat fields, causing systematic errors in the photometric calibration. We correct for this stray light in a two-step process. First, these runs are processed assuming that there is no stray light. Significant photometric residuals from comparisons with our PT calibration patches (§ 4.5.2) are attributed to the presence of the stray light. The peak-to-peak variation of this effect is 10%. We calculate a correction vector to compensate for this and apply it during a second processing of the runs. Once this is done, the data show no detectable residual effect from the scattered light.

#### 4.5.4. Assessment of Photometric Calibration

The primary sources of error are uncorrected extinction variations that affect both the PT patches and the 2.5 m scans, incidents of bad seeing in the 2.5 m scans coinciding with the crossing of PT patches, and statistical errors due to the small number of matching stars in certain PT patches in areas of the sky with low star density. Consequently, after final calibration, the data are subjected to a suite of tests to assess the reproducibility of photometric calibrations. These tests are neither exhaustive nor precise, but they do catch egregious errors. They are designed to uncover systematic calibration errors in the scan direction and across the imaging camera.

The principal tests for photometric accuracy include the following:

1. Comparison of photometry of objects in overlaps between adjacent runs. Figure 11 contains a typical set of plots, for the difference in our calibrated magnitudes for objects measured in the overlap region in one camera column (`camCol = 3`) for one pair of runs (752 and 756) for the five filters.

2. Calculation of fiducial points in the color-color diagrams of stars. Stars lie for the most part on a narrow one-dimensional locus in SDSS color-color space (Newberg & Yanny 1997; Fan 1999; Finlator et al. 2000). This locus is not straight but shows some bends, in particular between K and M stars in the plot of  $g^*-r^*$  versus  $r^*-i^*$ . We measure the location of this break in color to high accuracy with the stars in ten-frame intervals and look for trends in each run as a function of frame number. Figure 12 shows that the location of the stellar locus is stable in the EDR, with systematic drifts less than 0.01 mag. The *ri position* and *gr position* are the location of the bend in the stellar locus in the  $r^*-i^*$  versus  $g^*-r^*$  color-color diagram, (Fig. 13, top right). The *iz position* is the mean  $i^*-z^*$  for stars near that bend.

3. Comparison of calibration zero points with the mean sky brightness of a given run. Although the sky brightness varies in any given band as a function of time, it is close to constant at a given time across the camera. Thus, variations of the *measured* sky brightness as a function of camera column are an indication of photometric zero-point errors.

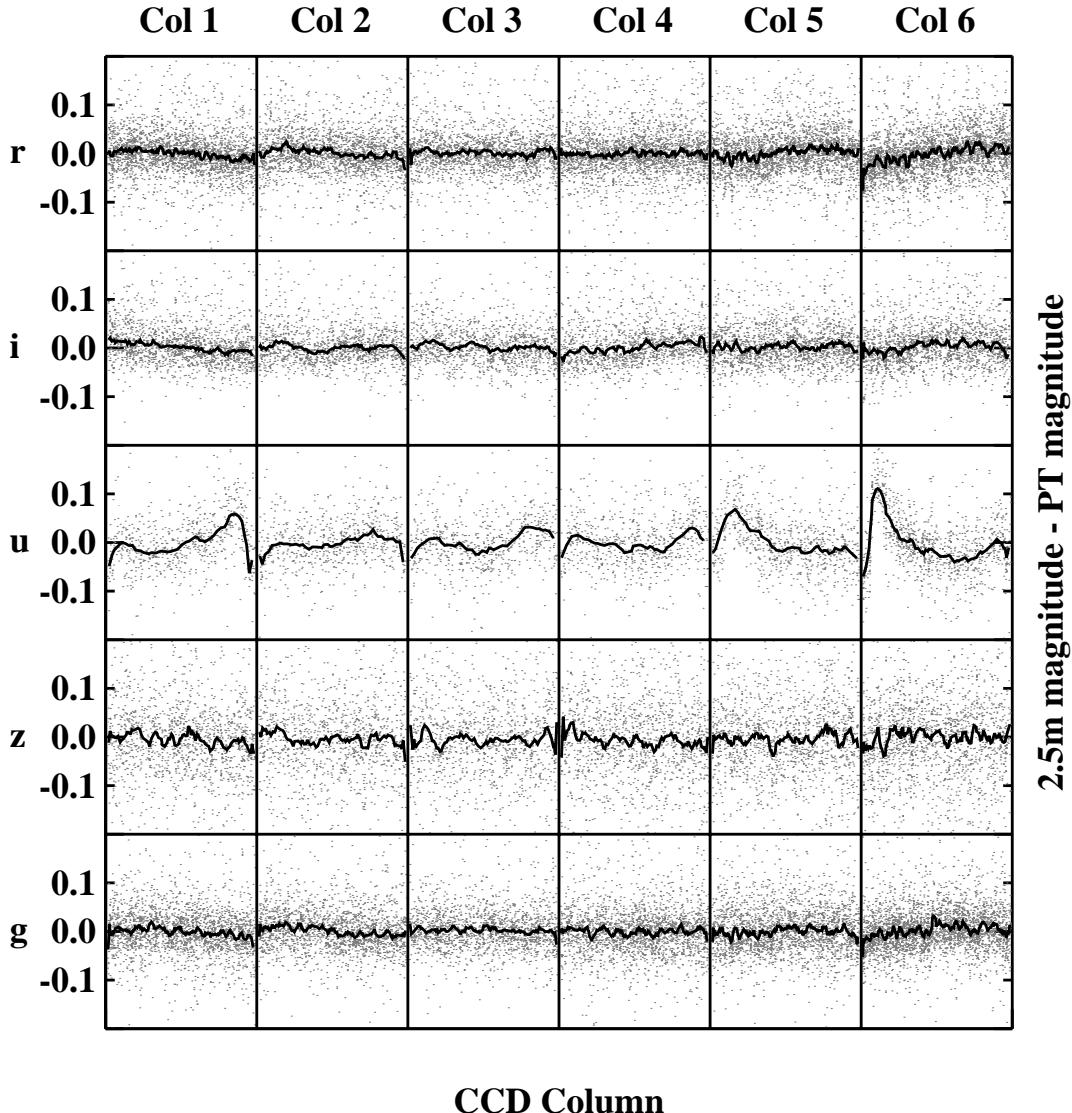


FIG. 10.—Zero-point deviations as a function of 2.5 m CCD column. This plot contains data from comparisons of 22,972 stars from runs 752 and 756 between the PT and the 2.5 m telescope and shows the 2.5 m magnitude minus the PT magnitude (which is equal to the zero point determined from each star). Scattered light creates deviations as great as 10% in the  $u$  band in columns 1 and 6.

4. Examination of systematics in the photometry of stars as measured by the 2.5 m and the PT in regions of overlap (see Fig. 10).

These tests consistently indicate that the photometric zero points for the EDR data are internally consistent to within  $\pm 2\%$  in all bands for 90% of the frames in the EDR data. The worst outliers are all in  $u$  and are up to 10% peak to peak in the worst of our data; as we saw above, these are due to ghosting in the  $u$  chip.

Because errors in different bands are often correlated, the colors of objects have smaller errors than might otherwise be indicated; stellar loci from different columns are typically aligned to better than 1% (measured in the  $griz$  bands). The median stellar loci from different runs are also aligned to better than 1%.

We discussed in § 4.4.5 that the estimated PSF magnitude errors are accurate to 10%–20%. The errors themselves are impressively small, as is manifest by the width of the stellar locus; the rms scatter in bright-star PSF photometry (where photon statistics are negligible), after

correcting for zero-point offsets, is 0.02 mag in  $g^*$ ,  $r^*$ , and  $i^*$ , and 0.03 mag in  $u^*$  and  $z^*$ . The separation between the stellar locus and quasars in Figure 13 is due to this excellent photometry.

In addition to errors on the internal consistency of the photometry, the tie of the 2.5 m photometry to an AB system has additional errors:

1. The transfer of zero points from the USNO to the 2.5 m currently does not use any color terms, thereby introducing errors in two places (USNO to PT and PT to 2.5 m, as described above). The error so introduced is of order 5%, although it may be somewhat smaller in  $r$ ,  $i$ , and  $z$ .

2. AB systems are based on absolute spectrophotometric measurements of standard stars and knowledge of the broadband response functions, as well as the overall spectrophotometric scale tied to Vega; these factors together produce uncertainties at the level of 5%.

3. The zero points given in Fukugita et al. (1996) ignore atmospheric absorption in the bandpass shapes, introducing an error of order 4% in  $u$  and 1% in  $g$ .

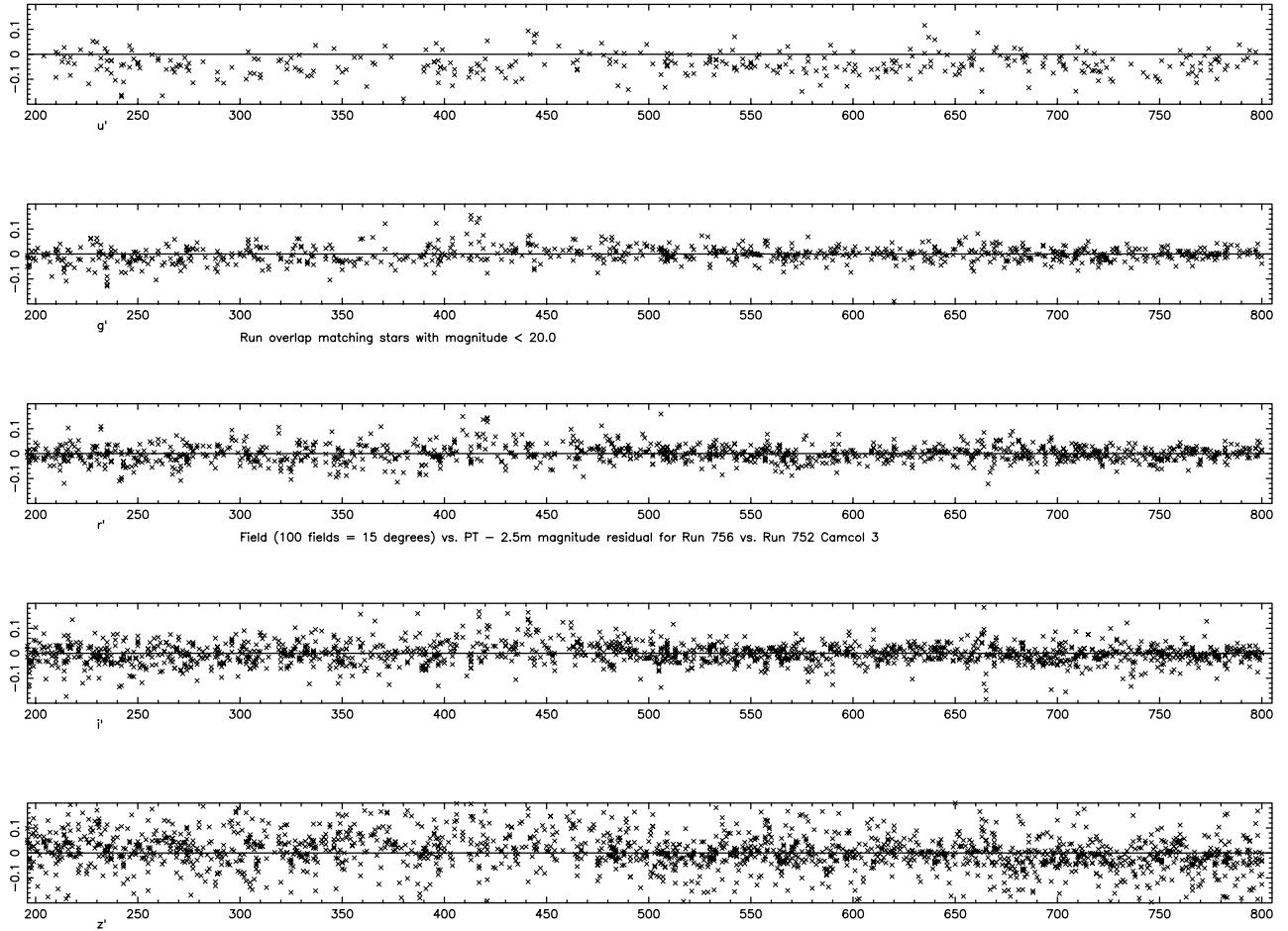


FIG. 11.—Difference of magnitude as a function of field number for objects that are measured in each of runs 752 and 756 in the overlap region of camera column 3. We inspect the complete set of these plots, six camera columns for each pair of overlapping runs. This set is typical, showing rms scatter of 0.02 mag in  $g^*$ ,  $r^*$ , and  $t^*$  and 0.03 mag in  $u^*$  and  $z^*$ .

Thus, the system zero points could differ from the AB zero points by as much as 10%.

#### 4.6. Imaging Data Quality Assessment

The quality of each field of data from the 2.5 m imaging camera is assessed based on a number of criteria. A quality is assigned for each criterion, and an overall quality is assigned based on the worst value of these. The possible values for quality in the class `Field` (Table 16) are

`AR_DQ_GOOD`: Data meet requirements.

`AR_DQ_ACCEPTABLE`: Data do not meet requirements but are usable for target selection; replace with new data if time permits.

`AR_DQ_HOLE`: Data do not meet requirements; the area of sky is small enough to not warrant replacing with new data.

`AR_DQ_BAD`: Data do not meet requirements.

`AR_DQ_MISSING`: Data are missing; the area of sky is small enough to not warrant obtaining new data.

For the EDR, all fields are in one of the first three categories. The quality for each field is indicated in Figure 8.

A set of flags is maintained for each field in which the quality for each of the criteria is recorded. These flags are 32 bit masks; there are four masks, one each for `GOOD`, `ACCEPTABLE`, `HOLE`, and `BAD`, with bits assigned for each cri-

terion used to judge quality. For example, if the seeing is judged acceptable, then the bit assigned to seeing in the “acceptable” mask is set. In addition, an overall quality is stored separately.

The following are the criteria used to assess data quality:

**PSF (seeing).**—The median point-spread function FWHM over all 30 (photometric) CCDs must be less than  $1.^{\circ}5$  for `GOOD` data. Because seeing is a function of wavelength, all measures of FWHM are corrected to the value that would be measured in the  $r$  band by scaling as  $\lambda^{-0.2}$ . The actual algorithm also smooths over time, so the quality of a field depends on more than just the instantaneous seeing. If the seeing fluctuates rapidly from field to field, quality is set to `AR_DQ_HOLE`. In addition, for fields where the FWHM is greater than  $1.^{\circ}5$  and the largest FWHM for any CCD is less than  $1.^{\circ}8$ , the quality is `AR_DQ_ACCEPTABLE`.

**Tracking.**—The telescope must track to within  $8''$  in the direction perpendicular to nominal path for `GOOD` data and  $18''$  for `ACCEPTABLE` data. This is necessary to not leave gaps between adjacent scan lines.

**PSF model.**—The postage-stamp pipeline (§ 4.3) outputs a status flag that indicates how accurately it was able to model the variable PSF in a field. This flag is converted to a quality of `GOOD`, `ACCEPTABLE`, or `HOLE`.

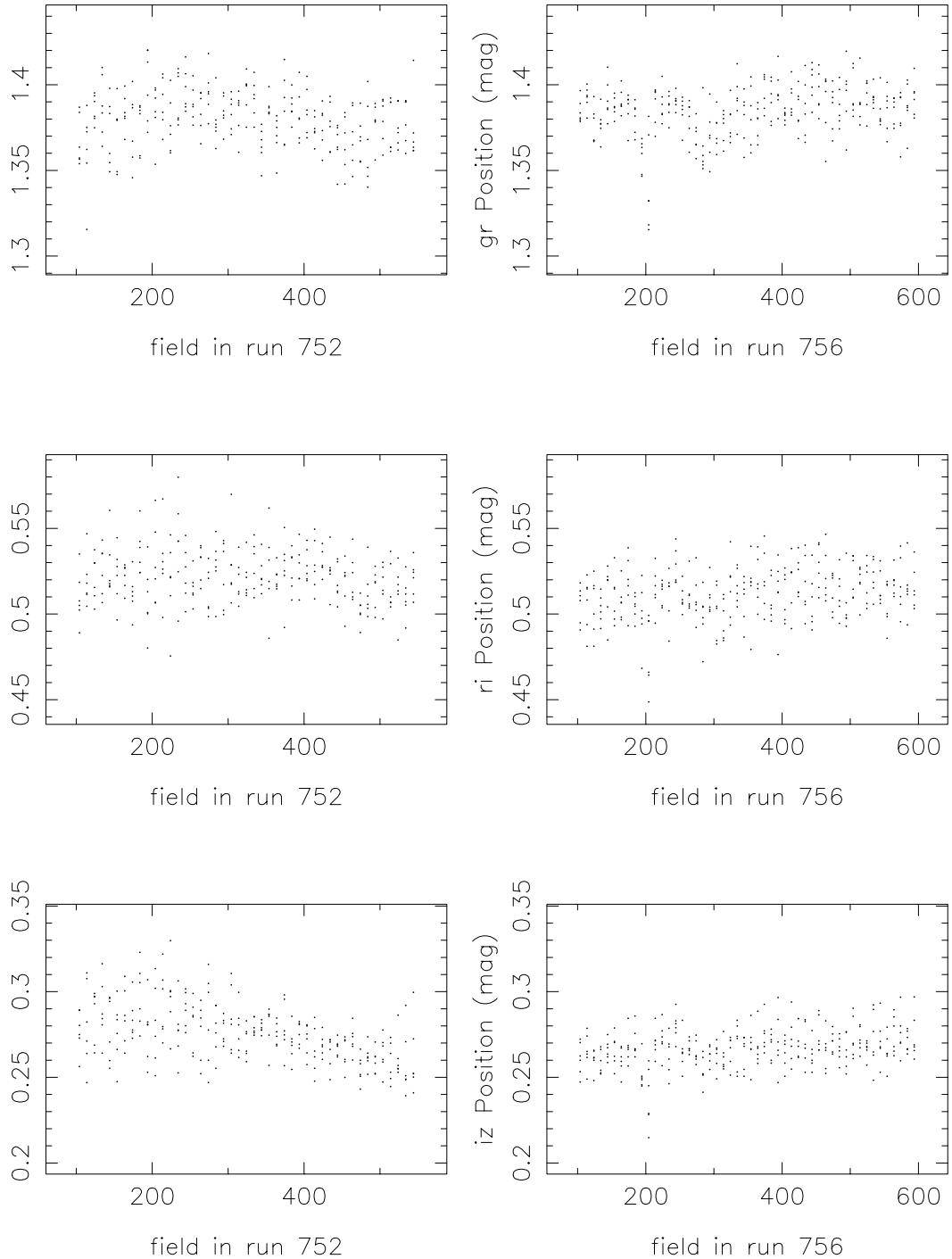


FIG. 12.—Location of the break in the stellar locus between K and M stars in the color-color diagrams, for runs 752 and 756 in the EDR

*Time-out.*—For some fields, the frames pipeline requires an inordinate amount of time to process, particularly if there are complex objects such as star clusters. In such cases the pipeline processing times out and no object file is created. For such fields, a HOLE is created.

*Photometricity.*—Data must meet certain conditions on photometry. At present, this quality assignment is not automated, because of the complexity involved in creating a robust test. Nonphotometric data are excised by hand.

*Manual.*—In some cases, it is desirable to override the overall automated quality assignment.

The parameter `badMask` in Table 16 encodes reasons why data were not declared GOOD.

The SKY and AMP bits are not used for the EDR. They are used to mask imaging data because of high sky values or bad amplifier noise, respectively.

A substantial portion of fields in the EDR have had quality set manually, overriding the PSF or tracking qual-

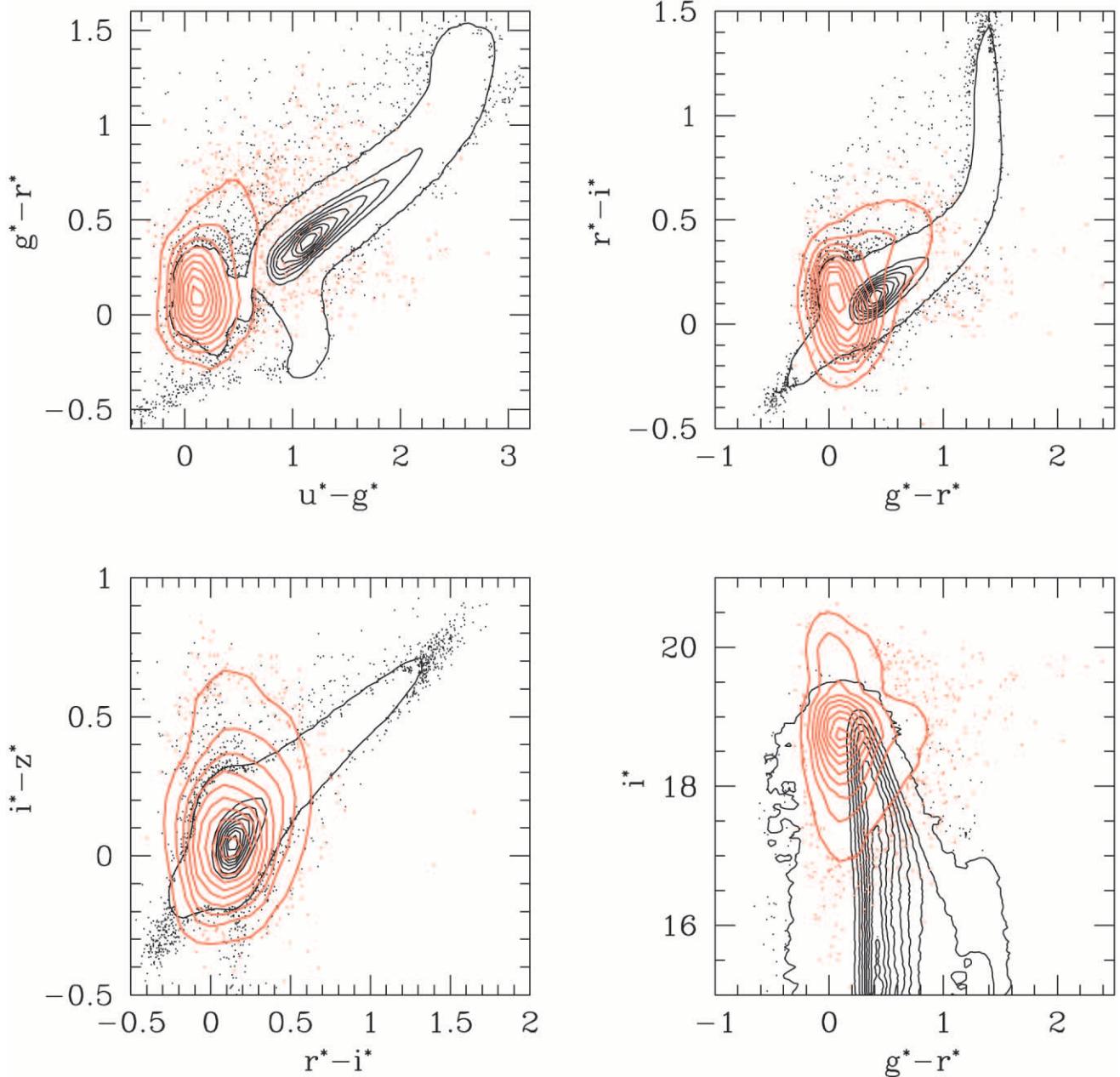


FIG. 13.—Color-color plots of quasars and stars as measured in the SDSS colors. Black contours and black dots are stellar sources. Red contours and red dots show the distribution of quasars in the EDR data.

ity in most cases. A list of all fields that have had the quality overridden are listed in Table 25. They have the AR\_FQ MANUAL flag bit set.

#### 4.7. Resolving Multiple Detections

For each stripe, 12 nonoverlapping but contiguous scan lines are defined parallel to the stripe great circle. Six of these scan lines are covered when the “north” stripe of the stripe is scanned, and the remaining six are covered by the “south” stripe. Each scan line is  $0^{\circ}20977$  wide and overlaps its neighbor by  $\approx 1'$ . Each overlap region is bisected by a line of constant  $\nu$ . Objects in overlap regions can have a detection in each of the scan lines, and we assign one of the scan lines to have the *primary detection* based on which side of the bisecting line it falls. Many queries want to consider only these objects, so we have constructed the Primary class in the Catalog Archive Server. The classes Star, Galaxy, Sky,

and Unknown are inherited from Primary. Other object detections are (usually) in the Secondary class, with its inherited classes StarSecondary, GalaxySecondary, Sky Secondary, and UnknownSecondary. Two kinds of objects are not in the Primary or Secondary class, and we put them in the Family class: objects that are tagged as BRIGHT (Table 9), and parent objects that have been successfully deblended.

Object detections in the area that overlaps adjacent stripes that we have not yet observed are in the Secondary class. These may be recovered by using the status flag in Table 6. The flag bits AR\_STATUS\_OK\_SCANLINE and AR\_STATUS\_OK\_STRIPE of status are set for all objects that are within one stripe.

Objects that lie close to the bisector between scan lines present some difficulty. Errors in the centroids or astrometric calibrations can place such objects on either side of the

TABLE 25  
EARLY DATA RELEASE MANUAL QUALITY

Run Number	Camera Column	Fields	New Quality
94.....	1–6	104–544	ACCEPTABLE
125.....	1–6	11–451	ACCEPTABLE
	1–6	11–300	ACCEPTABLE
	4	339	HOLE
		20	BAD
		461–480	ACCEPTABLE
		481–483	HOLE
752.....	1–6	484–500	ACCEPTABLE
756.....	1–6	732–802	ACCEPTABLE
		20	BAD
		21–60	ACCEPTABLE
1356.....	1–6	91–95	ACCEPTABLE
	1–6	19	ACCEPTABLE
	3	20	HOLE
1359.....	1–6	20–80	ACCEPTABLE

bisector. A resolution is performed at all bisectors, and if two objects that are a *primary detection* lie within 2" of each other, then one object is declared primary and the other secondary.

In the EDR, most data are taken along stripes that are isolated. Runs 752 and 756 cover stripe 10, and we plan to observe the stripes adjacent to it. Therefore, only objects that do not fall in an area that will overlap with the adjacent stripes may be marked as Primary detections. Runs 94 and 125 cover stripe 82, in the southern Galactic hemisphere. We do not plan to observe the stripes adjacent to stripe 82, so all objects may be marked as Primary detections.

Objects must satisfy other criteria as well to be labeled primary; an object must not be flagged BRIGHT (as there is a duplicate "regular" detection of the same object; § 4.4.2) and must not be a deblended parent (as the children are already included); thus, it must not be flagged BLENDED unless the NODEBLEND flag is set. These are put in the Family class.

The set of all primary detections from interleaving strips constitute a complete object catalog.

#### 4.8. Target Selection

The spectroscopic survey aims to observe complete samples of three categories of astronomical objects (Table 26):

Galaxies, selected to be brighter than  $r^* = 17.77$ , using Petrosian magnitudes;

Luminous red galaxies, selected by their color and magnitude to be a roughly volume-limited sample of the most luminous red elliptical galaxies to redshift  $\approx 0.45$ ; and

Quasars, selected by their distinctive colors in the SDSS photometric system, and by radio detection in the FIRST survey catalog.

The following sections describe the algorithms used to select targets for spectroscopic observation. The quantity prim Target records the results of target selection for each object by using the target bit masks in Table 27. An object can be targeted by more than one algorithm. The quasar and galaxy targets are marked as "tiled" in the table to indicate that we intend to obtain a uniform and complete sample. The TARGET\_STAR\_BROWN\_DWARF and TARGET\_HOT\_STD targets are also tiled, as described below.

TABLE 26  
BREAKDOWN OF FIBER ALLOCATION  
FOR A TYPICAL PLATE

Target	Number of Fibers
Total.....	640
Sky.....	32
Spectrophotometric standards.....	8
Reddening standards.....	8
Hot subdwarf standards.....	$\approx 2$
Galaxies:	
Main galaxy sample.....	$\approx 420$
LRG sample.....	$\approx 60$
Total.....	$\approx 480$
Quasar candidates.....	$\approx 80$
Unreserved:	
ROSAT All-Sky Survey sources .....	$\approx 5$
Stars.....	$\approx 15$
Serendipity objects .....	$\approx 7\text{--}8$
Quality-assurance targets .....	$\approx 2\text{--}3$
Total.....	$\approx 30^\circ$

The data included in the EDR are commissioning data and were used, among other things, to refine the algorithms discussed here. Table 20 lists the versions of the target selection pipeline that were used to select which objects to observe spectroscopically. After we observed the spectra, we ran a final version of target selection on all imaging objects. These results are stored in the PhotoObj class, while the results of the original target selection are stored in the SpecObj class, as summarized in Table 28. The exact algorithms used are not uniform throughout all of the data released, and Table 29 summarizes the important differences. The results of target selection are recorded in the quantity PrimTarget, which appears in two classes. The class SpecObj records the results of target that we used to select objects for spectroscopic observation (versions 2.2a, 2.5, and 2.7), while the class PhotoObj records the results of target run on the final photometric catalog (target version 2.13.8).

The galaxy magnitude limit has changed slightly, as has the criterion for selecting low surface brightness galaxies, and a number of changes have been made to quasar target selection. Since the pipeline software versions, the photometric calibration, and target selection criteria change, there will be slight mismatches between the value of PrimTarget for the same object recorded in class SpecObj and PhotoObj.

The quasar selection algorithm is substantially different in target version 2.13.8. A complete description of these improvements is beyond the scope of this paper.

##### 4.8.1. Galaxies

Images of galaxies are distinguished from images of stars by morphology. As described above (§ 4.4.6), we separate stars from galaxies using the difference between the PSF and model magnitude in  $r^*$ . Galaxy target selection requires a difference greater than 0.3 mag. During commissioning, we changed the magnitude limit of the sample several times (Table 29) and experimented with a fuzzy limit, whereby the probability that a galaxy would be selected was a declining function of  $r^*$  over a range of 0.1 mag. For the EDR, a sharp magnitude cut of  $r_p^* = 17.5$  yields a complete sample

TABLE 27  
TARGET SELECTION FLAGS

Name	Hexadecimal Bit	Tiled Target?	Description
Primary Targets			
TARGET_QSO_HIZ.....	0x1	Yes	High-redshift QSO
TARGET_QSO_CAP.....	0x2	Yes	Quasar at high Galactic latitude
TARGET_QSO_SKIRT.....	0x4	Yes	Quasar at low Galactic latitude
TARGET_QSO_FIRST_CAP.....	0x8	Yes	“Stellar” FIRST source at high Galactic latitude
TARGET_QSO_FIRST_SKIRT.....	0x10	Yes	“Stellar” FIRST source at low Galactic latitude
TARGET_QSO_FAINT.....	0x2000000	No	Stellar outlier; too faint to target
TARGET_QSO_REJECT.....	0x20000000	No	Stellar outlier, in excluded region
TARGET_GALAXY_RED.....	0x20	Yes	LRG
TARGET_GALAXY_RED_II.....	0x4000000	Yes	Cut II LRG
TARGET_GALAXY.....	0x40	Yes	Galaxy
TARGET_GALAXY_BIG.....	0x80	Yes	Large, relatively low surface brightness galaxy
TARGET_GALAXY_BRIGHT_CORE.....	0x100	Yes	Low surface brightness galaxy with bright fiber magnitude
TARGET_ROSAT_A.....	0x200	No	Also radio source
TARGET_ROSAT_B.....	0x400	No	AGN colors
TARGET_ROSAT_C.....	0x800	No	Bright or blue
TARGET_ROSAT_D.....	0x1000	No	Low-priority <i>ROSAT</i> target
TARGET_ROSAT_E.....	0x8000000	No	Too faint
TARGET_STAR_BHB.....	0x2000	No	Blue horizontal-branch star
TARGET_STAR CARBON.....	0x4000	No	Carbon star
TARGET_STAR_BROWN_DWARF.....	0x8000	Yes	Brown dwarf
TARGET_STAR_SUB_DWARF.....	0x10000	No	Subdwarf
TARGET_STAR_CATY_VAR.....	0x20000	No	Cataclysmic variable
TARGET_STAR_RED_DWARF.....	0x40000	No	Red dwarf
TARGET_STAR_WHITE_DWARF.....	0x80000	No	White dwarf
TARGET_STAR_PN.....	0x10000000	No	Planetary nebula
TARGET_SERENDIP_BLUE.....	0x100000	No	
TARGET_SERENDIP_FIRST.....	0x200000	No	
TARGET_SERENDIP_RED.....	0x400000	No	
TARGET_SERENDIP_DISTANT.....	0x800000	No	
TARGET_SERENDIP_MANUAL.....	0x1000000	No	
Secondary Targets			
TARGET_LIGHT_TRAP.....	0x1	No	Light trap
TARGET_REDDEN_STD.....	0x2	No	Reddening standard
TARGET_TEST_TARGET.....	0x4	No	Test target
TARGET_QA.....	0x8	No	Quality assurance target
TARGET_SKY.....	0x10	No	Blank sky
TARGET_SPECTROPHOTO_STD.....	0x20	No	Spectrophotometric standard
TARGET_GUIDE_STAR.....	0x40	No	Guide star
TARGET_BUNDLE_HOLE.....	0x80	No	Hole for a fiber bundle
TARGET_QUALITY_HOLE.....	0x100	No	
TARGET_HOT_STD.....	0x200	Yes	Hot subdwarf standard star

of galaxies, with a surface density on the order of 90 galaxies per square degree (see the discussion in § 4.9).

All magnitudes are corrected for extinction following Schlegel et al. (1998). Objects with fiber magnitudes brighter than 15 in  $g^*$  or  $r^*$ , or brighter than 14.5 in  $i^*$ , are rejected, to avoid problems with fiber cross talk in the spectrographs, for target version 2.7, runs 94 and 125, as indicated in Table 29. We also reject objects with SATUR set, which unfortunately rejects galaxies blended with saturated stars. Finally, the more recent versions of target reject objects with  $\text{PetroR50} < 2''$ , brighter than  $r^* = 15.5$  (version 2.7) and 15.0 (version 2.13.8), to cut down on contamination from very bright stars.

We have found that galaxy targets of very low surface brightness are often not real objects but are due to very low level scattered light within the camera. We therefore put in an explicit surface brightness cut. A *Petrosian half-light sur-*

*face brightness* is defined as the mean surface brightness within the radius,  $\text{PetroR50}$ , that includes half the Petrosian flux. All objects with Petrosian surface brightness fainter than the value listed in Table 29 are not targeted (the surface brightness limit of 30.0 for version 2.5 indicates that, in effect, no surface brightness cut was applied).

For target version 2.13.8, we check that for faint objects, the local sky value in  $r$  is sensible. Objects with surface brightness between 23.0 and 24.5 are only selected if the value of the sky determined locally (i.e., on scales of  $\sim 2'$ , sky in class *PhotoObj*) and that on a frame (10' scales, sky in class *Field*) are within 0.05  $\text{asinh}$  magnitude per square arcsecond of one another; the local sky of a child includes the contribution of its siblings, so this cut discriminates against pieces of fluff inadvertently deblended from bright galaxies. These objects are flagged both TARGET\_GALAXY and TARGET\_GALAXY\_BIG.

TABLE 28  
TARGET SELECTION VERSIONS: RUNS AND PLATES

target Version	Runs	Plates <sup>a</sup>	primTarget Class
v2.2a .....	752, 756	266–315, 363	SpecObj
v2.5 .....	1336, 1339, 1356, 1359	349–362, 366–367	SpecObj
v2.7 .....	94, 125	383–416	SpecObj
v2.13.8 ...	All	All	PhotoObj

<sup>a</sup> Plate 324 is dominated by a special selection of targets in the stellar locus.

Finally, galaxies with fiber magnitudes brighter than  $r^* = 19$  that do not otherwise make the surface brightness cut are targeted. These are flagged TARGET\_GALAXY\_BRIGHT\_CORE. Such objects are quite rare. The resulting sample is contaminated at the  $\approx 1\%$  level by close double stars that are not separated by the deblender.

#### 4.8.2. Luminous Red Galaxies

It has long been known that the most luminous galaxies in clusters are a very homogeneous population (e.g., Postman & Lauer 1995 and references therein); they have a very narrow range of color and intrinsic luminosity. Because these objects are intrinsically very luminous, they can be observed to great distance.

We therefore selected such objects by a variant of the method of photometric redshifts (Connolly et al. 1995a), with the aim of selecting a uniform, approximately volume-limited sample of objects with the reddest colors in the rest frame (Eisenstein et al. 2001). We term these objects “luminous red galaxies,” or LRGs. It turns out that this can be simply done with cuts in the  $(g^*-r^*, r^*-i^*, r^*)$  color-

color-magnitude cube. Because the 4000 Å break moves from the  $g$  band to the  $r$  band at a redshift  $z \approx 0.4$ , two different sets of selection criteria are used, which we term cut I and cut II. First, we rotate to a color basis that is aligned with the galaxy locus in the  $(g^*-r^*, r^*-i^*)$ -plane as follows:

$$\begin{aligned} c_{\perp} &= (r^* - i^*) + \frac{1}{4}(g^* - r^*) + 0.18, \\ c_{\parallel} &= 0.7(g^* - r^*) + 1.2[(r^* - i^*) - 0.18]. \end{aligned} \quad (16)$$

All colors are measured using model magnitudes, and all quantities are corrected for Galactic reddening following Schlegel et al. (1998).

Cut I and cut II objects used star-galaxy separation criteria that changed from one chunk to another (Table 29). For cut I, the flux and color cuts are

$$\begin{aligned} r_p^* < 19.2, & \text{ roughly 1.5 mag fainter than main galaxies ;} \\ \text{SB}_{P,r^*} < 24.2, & \text{ roughly the same as main galaxies ;} \\ r_p^* < 13.1 + c_{\parallel}/0.3, & \text{ luminosity cut ;} \\ |c_{\perp}| < 0.2, & \text{ color cut .} \end{aligned} \quad (17)$$

This yields an approximately volume-limited sample of galaxies to  $z = 0.37$  with additional galaxies up to  $z \approx 0.45$ . These objects are flagged GALAXY\_RED.

To reach higher redshifts, we use cut II:

$$\begin{aligned} r_p^* < 19.5, & \text{ a little fainter than above ;} \\ \text{SB}_{P,r^*} < 23.3 \text{ or } 24.2, & \\ c_{\perp} > 0.45 - \frac{1}{6}(g^* - r^*), & \text{ color cut ;} \\ g^* - r^* > b_{g^*-r^*} + \frac{1}{4}(r^* - i^*), & \text{ color cut ;} \end{aligned} \quad (18)$$

This selects red galaxies at  $0.42 < z < 0.55$ . The surface brightness cut, and the value of the  $g^*-r^*$  color-cut offset,

TABLE 29  
TARGET SELECTION VERSIONS: SELECTION CRITERIA

Criterion	v2.2a	v2.5	v2.7	v2.13.8
Main Galaxy Sample				
petroMag limit .....	17.65	17.62	17.67	17.77
Fuzz .....	0.10	0.00	0.00	0.00
SB limit .....	23.5	30.0	23.5	24.5
Bright limit ( $g^*r^*i^*$ ) .....	None	None	(15.0, 15.0, 14.5)	(15.0, 15.0, 14.5)
Small bright limit .....	None	None	15.5	15.0
Check local sky? .....	No	No	No	Yes
Luminous Red Galaxy Sample				
Galaxy morphology (I) ....	objc_type == 3	objc_type == 3	$r_{\text{psf}}^* - r_{\text{model}}^* > 0.3$	$r_{\text{psf}}^* - r_{\text{model}}^* > 0.3$
Galaxy morphology (II) ....	objc_type == 3	objc_type == 3	$r_{\text{psf}}^* - r_{\text{model}}^* > 0.3$	$r_{\text{psf}}^* - r_{\text{model}}^* > 0.5$
$b_{g^*-r^*}$ (eq. [18]) .....	1.35	1.35	1.35	1.30
$r^*-i^*$ (II)? .....	No	No	No	Yes
SB limit (II) .....	23.3	23.3	23.3	24.2
Quasar Sample				
Bright $i^*$ limit .....	16.5	15.0	15.0	
Color rejection? <sup>a</sup> .....	Only A stars	Yes	Yes	
QSO_HIZ ( $u^*g^*r^*i^*$ )? .....	No	Yes <sup>b</sup>	Yes	

<sup>a</sup> The four color rejection criteria are hot white dwarfs (eq. [19]); unresolved M dwarf–white dwarf pairs (eq. [20]); A stars (eq. [21]); and blue objects, from the high-redshift quasar sample (eq. [22]).

<sup>b</sup> These high-redshift quasar candidates were flagged as QSO\_CAP or QSO\_SKIRT instead of QSO\_HIZ.

$b_{g^*-r^*}$ , was different in different versions of target selection, as indicated in Table 29. Despite the deeper flux cut, this sample is still flux limited rather than luminosity limited. These cut II objects are flagged both as GALAXY\_RED\_II and GALAXY\_RED.

The luminosity cut in cut I (eq. [17]) is a good approximation to a passively evolving old stellar population from  $z \approx 0.15$  to  $z \approx 0.45$ , but it is too permissive at lower redshift. At  $z < 0.20$ , one should not use the GALAXY\_RED flag to extract the low-redshift extension of the LRG volume-limited sample. Rather, one must use the redshift extracted from the main galaxy sample to construct the absolute magnitude and rest-frame color and select galaxies that fall in the region populated by the high-redshift LRGs. This is described in more detail in Eisenstein et al. (2001).

The resulting LRG sample has nearly constant comoving number density to  $z \approx 0.4$  with a passively evolved luminosity threshold that is close to constant. This suggests that the sample is indeed close to the volume-limited goal. At  $z > 0.4$ , the sample is effectively flux limited but still selects intrinsically red galaxies. The sample selection is also quite efficient: 97% of the selected targets are at  $z > 0.15$ .

The two LRG cuts together yield roughly 12 targets per square degree that are not already in the main galaxy target selection sample, typically 10 of which are cut I targets and two of which are cut II targets.

#### 4.8.3. Quasars

Images of quasars are characterized in the optical by high luminosity and very high surface brightness, making them morphologically indistinguishable from stars at all but very low redshifts ( $z \leq 0.6$ , at the resolution of the SDSS). Quasar spectra, however, are characterized by roughly power-law continua and strong emission lines; as a result, quasar spectra are quite distinct from the roughly blackbody spectra of stars.

The dominance of the continuum on the colors of quasars is severely reduced because of absorption by the Ly $\alpha$  forest and Lyman limit systems (at rest wavelengths below 1216 and 912 Å, respectively). For the SDSS filter set, absorption by the Ly $\alpha$  forest is significant for  $z \geq 2.2$ , and quasars become increasingly red with larger redshift (Fan 1999; Richards et al. 2001). Fortunately, this effect also gives quasars colors that are distinctly different from those of normal stars.

Models of quasars by Fan (1999) show that the quasar locus is indeed well separated from the stellar locus except at  $z = 2.7\text{--}2.8$ , where quasars have SDSS broadband colors that are essentially indistinguishable from early F and late A stars. Thus, the majority of quasars have colors quite distinct from those of stars, which, unlike quasars, lie on a nearly one-dimensional locus in  $(u^*-g^*, g^*-r^*, r^*-i^*, i^*-z^*)$  color-color-color-color space (Newberg & Yanny 1997; Fan 1999; Finlator et al. 2000). Using this property, we identify quasar candidates primarily as outliers from the stellar locus in this four-color space.

The quasar target selection algorithm first models the stellar locus, following Newberg & Yanny (1997), as an elongated tube with an elliptical cross section. In practice, this is done in two stages, once for the  $(u^*-g^*, g^*-r^*, r^*-i^*)$  color cube, and once for the  $(g^*-r^*, r^*-i^*, i^*-z^*)$  color cube. The PSF colors of each object are examined in

turn, and it is determined whether they are consistent with lying inside the stellar locus, incorporating the photometric errors. The quoted errors in PSF magnitudes from frames are used, after adding  $\approx 0.03$  mag errors in quadrature in each color to reflect the uncertainty in the photometric calibration. Quasar candidates are those objects that lie more than  $4\sigma$  from the stellar locus. Colors are measured using PSF magnitudes and are corrected for reddening using Schlegel et al. (1998) throughout.

The quantity objFlags in class photoObj uses the bit masks in Table 9 to indicate objects whose photometry (and therefore colors) may be in error. In particular, we reject objects flagged as BLENDED (although we accept the deblended children of blended parents), BRIGHT, SATURATED, or EDGE in any band. Objects detected as MOVING at a rate greater than 0.2 pixels per frame with high significance are rejected as asteroids (although this rejection is no longer needed in the current algorithm: asteroids have colors that put them cleanly in the stellar locus; Ivezić et al. 2001). Deblended children with PEAKCENTER, NOTCHECKED, or DEBLEND\_NOPEAK set in any band are signs of deblending problems; these objects are also rejected. Objects with INTERP\_CENTER set have a cosmic ray or bad column within 3 pixels of their center; empirically, we find that many false quasar candidates are found with  $i^* < 16.5$  and this flag set, so we reject all such objects. There are a few bad columns that are not properly interpolated over by the photometric pipeline, and so we reject objects in CCD columns 1019–1031 in camCol 5.

##### 4.8.3.1. Main Quasar Sample

At one point, we had considered separate criteria in regions of high and low stellar density, as we were concerned about changing stellar populations at low latitudes. For this reason, objects that pass the main quasar target selection criteria have the TARGET\_QSO\_CAP (referring to the Galactic cap) or TARGET\_QSO\_SKIRT (the ring around Galactic cap region) target mask bit set in primTarget of class specObj, depending on their Galactic coordinates and, in particular, the estimated stellar density from the Bahcall & Soneira (1980) model. However, we found that the effectiveness of the algorithm was independent of stellar density, so the selection of the two is identical.

An outlier from the  $(u^*-g^*, g^*-r^*, r^*-i^*, i^*-z^*)$  color cube is a quasar candidate if it passes the above gauntlet of tests and has  $15.0 < i^* < 19.0$ . The bright limit is set to avoid problems with cross talk in the spectrographs. As indicated in Table 29, the bright limit used for target version 2.2a, runs 752 and 756, was  $i^* = 16.5$ . The object need not be stellar if it lies blueward of the main galaxy locus; this allows low-redshift active galactic nuclei (AGNs), whose host galaxy is visible, to be selected.

We reject some quasar candidates if they have colors that match known types of stars that lie outside the main stellar locus. In particular, hot white dwarfs are excluded when

$$\begin{aligned} -0.8 < g^* - r^* < -0.2 \quad \text{and} \quad -0.6 < r^* - i^* < -0.2 \\ \text{and} \quad -1.0 < i^* - z^* < 0.0 ; \end{aligned} \quad (19)$$

unresolved M dwarf–white dwarf pairs are excluded when

$$0.0 < g^* - r^* < 1.6 \quad \text{and} \quad 0.6 < r^* - i^* < 2.0 ; \quad (20)$$

and A stars are excluded when

$$0.9 < u^* - g^* < 1.5 \text{ and } -0.35 < g^* - r^* < 0.0. \quad (21)$$

The quasar locus crosses the stellar locus for quasars of  $z \sim 2.7$ , so we *include* target objects when

$$0.65 < u^* - g^* < 1.5 \text{ and } 0.0 < g^* - r^* < 0.2, \quad (22)$$

even if they are not outliers from the stellar locus.

Each of these three regimes of stars was rejected in the EDR (as indicated in Table 29) except with target version 2.2a, runs 752 and 756, when only the A star rejection was applied. Inclusion of objects in the region where the stellar and quasar loci cross was performed for all three versions of target used for the EDR.

The SDSS catalog is matched against the FIRST catalog of radio sources (Becker et al. 1995), and stellar objects with radio counterparts within  $2''$  and with  $i^* < 19.0$  are selected, using the mask bit TARGET\_QSO\_FIRST\_CAP or TARGET\_QSO\_FIRST\_SKIRT, depending on their position on the sky.

#### 4.8.3.2. High-Redshift Quasar Sample

The mask bit in primTarget that we use for the high-redshift quasar sample is TARGET\_QSO\_HIZ. Outliers from the stellar locus in the  $(g^* - r^*, r^* - i^*, i^* - z^*)$  color cube with  $i^* < 20.0$  are selected, subject to the rejection criteria given above in equations (19)–(21). In addition, candidates must be classified as stellar, as they will have redshifts above  $z = 3.5$  for the most part (see the discussion in Fan et al. 2001b). For target versions 2.5 and 2.7, we also targeted objects that lie in the region of color-color space for  $z > 3$ , are still outliers in the  $(u^* - g^*, g^* - r^*, r^* - i^*)$  cube, and have  $u^* - g^* > 1.5$  and  $i^* < 20.0$  (see Table 29).

Although low-redshift quasars manifest themselves most dramatically as outliers in  $u^* - g^*$ , they are also outliers in the  $g^*r^*i^*z^*$  color cube and would thus dominate the (nominally) high-redshift color cut, which goes fainter than the main quasar sample. Therefore, we rejected very blue objects that satisfied

$$u^* - g^* < 0.9 \text{ and } g^* - r^* < 0.8 \text{ and } i^* > 19.0, \quad (23)$$

even if they were outliers in the  $g^*r^*i^*z^*$  color cube. Table 29 indicates that this rejection was not done in target version 2.2a.

#### 4.8.3.3. Faint, Reject, and Magnitude Outlier Objects

For target versions 2.2a, 2.5, and 2.7, objects that lie outside the stellar locus but are fainter than the magnitude limit are flagged as TARGET\_QSO\_FAINT; they are not targeted but are marked as being of potential interest. We emphasize, however, that the algorithm is not optimized fainter than  $i^* = 20.0$ , so the fraction of objects flagged as TARGET\_QSO\_FAINT that turn out to be quasars will decrease significantly with magnitude. Objects with this flag set will only have spectra if they were targeted by another module.

#### 4.8.3.4. Performance of Quasar Target Selection

Quasar candidates are selected primarily as objects with colors distinct from stars', allowing the serendipitous discovery of a variety of objects of unusual color, including those that do not conform to "standard" quasar colors. We have thus found a number of very unusual broad

absorption line quasars, extreme emission-line galaxies, high-redshift ( $z > 0.7$ ) starburst and post-starburst galaxies, and a handful of objects whose spectra defy classification.

This algorithm selects of order 15 objects per square degree. Our target efficiency requirement is that at least 65% of our quasar candidates should in fact be quasars. The EDR is very close to meeting this requirement, and we certainly expect future data releases to meet or exceed this requirement. The principal contaminants are compact blue star-forming galaxies at low redshift; A stars; white dwarfs and M dwarf–white dwarf pairs that the above color cuts do not completely remove; and a variety of exotica such as  $z > 0.4$  compact E+A galaxies, whose Balmer break gives them colors very similar to  $z = 4$  quasars. Comparison with quasars from the literature shows that our completeness is of order 90%, which is consistent with our completeness requirement. Known quasars that are not targeted are rejected for a wide variety of reasons: many are due to cosmetic defects, and many  $z = 3.5$  and  $z = 4.5$  quasars are missing because of problems with the quasar target selection code, which have been corrected since target was run to define the plates in the EDR.

#### 4.8.4. Other Science Targets

A variety of other science targets are selected. These include optical counterparts of *ROSAT* sources, stars, and serendipity targets. These samples are not complete (with the exception of spectrophotometric standards and brown dwarfs) but are assigned to excess fibers left over after the main samples of galaxies, LRGs, and quasars are tiled (§ 4.9).

##### 4.8.4.1. *ROSAT* Targets

For typical values of the ratio of optical to X-ray flux of various classes of Galactic and extragalactic sources, there is an excellent match between the depth of the SDSS in the optical and that of the *ROSAT* All-Sky Survey (RASS; Voges et al. 1999) in X-rays. Objects detected in SDSS imaging data are positionally matched against X-ray sources from the RASS catalogs. SDSS objects within the RASS positional error circles (commonly  $10''$ – $20''$  radius) are scrutinized further by target selection algorithms tuned to select likely optical counterparts to the X-ray sources. In order of decreasing priority, spectra are obtained for SDSS-RASS coincidences that (1) are also radio sources (flagged ROSAT\_A); (2) have SDSS optical colors of AGNs or quasars (ROSAT\_B); (3) fall in a broad intermediate-priority bin that includes stars and galaxies that are bright, moderately blue, or both (ROSAT\_C); or (4) are otherwise bright enough for follow-up spectroscopy (ROSAT\_D). Objects flagged ROSAT\_E are not targeted; they fall within the *ROSAT* error circle but are either too faint or too bright for our spectrographs. Objects with fiber magnitude brighter than 15 in  $g^*$ ,  $r^*$ , or  $i^*$  are not targeted, because of cross talk between fibers in the spectrograph.

This algorithm targets of order 10–15 objects per plate; over half are AGNs or quasars.

##### 4.8.4.2. Stars

A variety of stars are also targeted by their distinctive colors. Specific stellar spectroscopic target categories include blue horizontal-branch stars (flagged STAR\_BHB), carbon stars, both dwarf and giant (STAR\_CARBON), low-luminosity

subdwarfs (STAR\_SUB\_DWARF), the central stars of planetary nebulae (STAR\_PN), cataclysmic variables (CATY\_VAR), red and brown dwarfs (STAR\_RED\_DWARF and STAR\_BROWN\_DWARF), and hot white dwarfs (STAR\_WHITE\_DWARF). The planetary nebula and cataclysmic variable categories are essentially completely contained within the quasar target selection region of color-color space and are much rarer than quasars; only a handful of these objects have been found. The definition of red dwarfs was very inclusive, resulting in the observations of a large number of ordinary M dwarfs. The brown dwarfs are very rare (typically one every several plates); unlike the other categories of stars, these are tiled. However, in the EDR, the algorithm for selecting these objects was modified several times, making the formal definition of a complete sample difficult.

In addition, of order 30% of the objects targeted as quasars turn out, in fact, to be stars, for the most part outliers from the stellar locus by definition (but occasionally objects in the stellar locus as well, as a result of early errors in the code). This is of course not a complete sample, but it contains several thousand spectra of uniform data quality for stars of all spectral types.

#### 4.8.4.3. Serendipity Targets

Finally, we have an open category of targets, referred to as *serendipity*, whose criteria are subject to change as different regions of parameter space are explored. There are several categories of these in the EDR. They are flagged as SERENDIP\_RED, SERENDIP\_BLUE, and SERENDIP\_DISTANT for objects lying outside the stellar locus in color space (the latter refers to distance from the stellar locus). These overlap somewhat with the quasar and some of the stellar target selection categories. Objects coincident with FIRST radio sources are flagged SERENDIP\_FIRST; this matching goes fainter than the equivalent for quasar target selection and has no restriction on stellarity. Finally, SERENDIP\_MANUAL targets are selected by hand, as the name implies. A wide variety of interesting objects have emerged from serendipity spectroscopic targets, ranging from highly unusual categories of white dwarfs to radio-loud broad absorption line quasars (Menou et al. 2001).

#### 4.8.5. Calibration Targets

In addition to science targets, about 50 fibers out of each plate of 640 are assigned by target for calibration spectra. Since spectroscopy requires good sky subtraction, 32 of these calibration fibers are assigned to regions where the photometric pipeline detects no objects (in fact, the photometric pipeline outputs several such "sky objects" in each frame for this purpose). These sky fibers are distributed so that each bundle of 20 fibers contains one; this ensures that the sky fibers are reasonably uniformly distributed around the plate. In addition, 18 fibers per plate are assigned to standard stars of various sorts; eight are spectrophotometric standards, chosen to be F dwarfs, eight are reddening standards, chosen to be F subdwarfs, and two are hot subdwarfs (Table 26). The spectrophotometric standards are also used in correcting the strong telluric water absorption bands in the red end of the science-target spectra (the smoothness of the F subdwarf spectra makes them ideal for this purpose). Conversely, the hot subdwarfs are useful for calibrating the blue end of the science-target spectra. Finally, as their name

implies, the reddening standards will be used for determining the reddening of our Galaxy.

Further details of the spectroscopic calibration process are discussed in § 4.10.1.

#### 4.8.6. Target Culling

In regions of particularly poor seeing, the colors of stars were occasionally systematically off, especially when earlier versions of Photo were used. This caused the target selection algorithms that select outliers in color-color space to choose an inordinate number of targets. We corrected for this by manually culling targets from the list.

In particular, target candidates were culled from run 752 (target version 2.2a) in 61 fields because of problems caused by poor seeing. In fields with five or more quasar candidates, we removed candidates that were not also selected as galaxy targets, in the following categories:

TARGET\_QSO\_HIZ, TARGET\_QSO\_CAP, TARGET\_QSO\_SKIRT,  
 TARGET\_QSO\_FIRST\_CAP, TARGET\_QSO\_FIRST\_SKIRT,  
 TARGET\_SERENDIP\_BLUE, TARGET\_SERENDIP\_RED,  
 TARGET\_SERENDIP\_DISTANT, TARGET\_STAR\_SUB\_DWARF,  
 TARGET\_STAR\_WHITE\_DWARF, TARGET\_STAR\_CARBON.

Similarly, target candidates in runs 94 and 125 (target version 2.7) were culled from 52 fields. In addition, all QSO\_HIZ candidates in run 94 in camera column 6 were removed from the list. This was necessary because of a problem with the characterization of the point-spread function in the version of psp used at the time.

After the plates from runs 94 and 125 were drilled, we discovered that many quasar candidates were targeted, as a result of an error in the target selection code. Spectra were collected for these targets, but they may be removed from quasar analyses by rejecting objects that satisfy the criteria  $0.65 < u^* - g^* < 0.9$  and  $0.2 < g^* - r^* < 0.9$ .

#### 4.9. Plate Definition

Once target selection has been completed, the spectroscopic plates need to be positioned and fibers assigned to targets. As noted above, although each plate can accommodate 640 fibers, about 50 of these are reserved for calibration targets, leaving around 590 fibers for science targets; for plates 266–315 the available number was 588, and for later plates the number was 592. As described in this section and in Table 27, a number of object types are potential spectroscopic science targets. However, only for some of these do we attempt to obtain a uniform and complete sample, the so-called tiled targets, as indicated in Table 27. Other science targets, such as most types of stars, *ROSAT* sources, and serendipity targets, are assigned leftover fibers according to a set of priorities after the plates are positioned and fibers are assigned to tiled targets.

To position the plates, we could simply place them on a uniform grid, but given the clustering of the galaxy targets that dominate the sample, this procedure would give some plates substantially more targets than available fibers, while others would have large numbers of extra fibers. To make the best use of observing time, an algorithm based on cost minimization determines the optimal positioning of plate centers in order to maximize the number of tiled targets on which a fiber can be placed. This spaces plates more closely together in regions with a

larger density of galaxies, increasing the completeness and efficiency of the spectroscopic survey. This process of optimally placing spectroscopic plates is referred to as *tiling*.

Because of the finite diameter of the fibers and their cladding, fibers cannot be placed more closely than 55" from one another. For each group of targets within 55", we put fibers on the largest possible set of the highest priority objects; in case there is a tie, we decide on the competing configurations based on the next highest priority objects, and so on. Brown dwarf and hot standard star targets have priority over quasar targets, which have priority over LRGs and main-sample galaxies. The priorities of each object are stored in a parameter *priority*; a higher value means higher priority. This procedure means that if two objects of the same priority are within 55", one is rejected at random. Roughly 8%–9% of selected objects are not observed for this reason, in regions covered by a single plate. In regions of plate overlap, of course, both objects in such a close pair can be targeted; about 30% of the sky is covered by such overlaps. Finally, there is another constraint, which is that no fiber can be placed within 100" of the center of a plate; this excludes a very small number of targets.

As a consequence of an error in an early version of the tiling pipeline, for the earliest plates (266–315) three gaps were left in the assignment of fibers to tiled targets. Because of the nature of the bug, these occurred in regions where plates overlapped; also, they unfortunately tended to occur in the densest regions. They occurred in the overlaps between plates 270 and 271, plates 312 and 313, and plates 314 and 315.

The bookkeeping of the tiling results has not been provided in this data release, including whether a fiber was eliminated because of a collision and whether it was originally targeted in the “drilling” version of the data reductions. Future data releases will include this information. However, it is possible to evaluate the completeness from the data themselves by calculating the fraction of each target type (as determined in the released version of the reductions) that have spectra; simply be careful to include the effect of fiber collisions. We have determined that for the south equatorial stripe and the *SIRTF* First Look field, the effective flux limit for the main galaxy sample is about  $r^* = 17.5$ ; for the north equatorial stripe, the effective flux limit is about  $r^* = 17.6$ . Finally, because there may be incompleteness in regions covered by one observed plate and one unobserved plate, it is important to know the locations of the unobserved plates when evaluating the window function of the survey. Thus, we include in Table 4 the locations of plates not released here.

As discussed above in § 4.8.6, we found that target selection did poorly in a few regions of sky of particularly bad seeing, especially with earlier versions of the imaging pipelines. The symptom was typically an excessive number of quasar candidates due to poor photometry. We were forced in a few regions therefore to cull the spectroscopic targets.

#### 4.10. Spectroscopic Pipelines

The spectroscopic pipelines, *spectro2d* and *spectro1d*, reduce the two-dimensional spectrograms produced by the spectrographs to flux- and wavelength-calibrated spectra and then measure emission and absorption features, classify the spectra, and measure redshifts.

The *spectro2d* pipeline reduces the raw data and calibration images from the red and blue CCD cameras from each spectrograph and produces merged, co-added, calibrated spectra, noise estimates, and mask arrays (see Table 11) for analysis by the *spectro1d* pipeline. All wavelengths are expressed in angstroms and are *vacuum* wavelengths corrected to the heliocentric frame, while the flux density  $f_\lambda$  is in units of  $10^{-17} \text{ ergs cm}^{-2} \text{ s}^{-1} \text{ \AA}^{-1}$ .

The *spectro1d* pipeline determines emission and absorption redshifts, classifies spectra by object type, and measures lines in each spectrum. The class *specObj* (Table 12) contains parameters measured for the entire spectrum, as well as links to the *Plate* that the spectrum is from, the *PhotoObj* information for the object, and links to objects in Table 13: *SpecLine*, lines identified and measured; *SpecLineName*, names for the lines; *SpecLineIndex*, equivalent widths and redshifts; *CrossCorrelationRedshift*, measures of the cross-correlation redshift; and *EmissionRedshift*, measures of the emission redshifts.

We also provide the flux- and wavelength-calibrated spectrum, the spectrum with continuum subtracted, and the estimate of the  $1\sigma$  error per pixel, all in the units above, plus the mask array, as described in § 2.5.

##### 4.10.1. Extraction and Calibration of Spectra

The *spectro2d* pipeline works in two stages. First, it reduces the data from each chip from each 15 minute exposure separately, and then it combines the results from each exposure.

After bias subtraction, each raw image is divided by a uniformly illuminated flat to take out pixel-to-pixel variations. The flat-field spectra (i.e., the flat field observed through the fibers) are then traced on the CCD: for each fiber, the flat-field image centroid in column position is fitted by a polynomial in row number. The flat-field spectra are optimally extracted, assuming a Gaussian profile for each one, simultaneously fitting a low-order polynomial to scattered light. These fits will also be used for the first-order object extraction of the science and arc frames, allowing for small offsets. The arc lamp spectra are extracted as well, and centroids of the lines are measured, to which we fit a fifth-order Legendre polynomial.

The flat-field spectra are then wavelength-calibrated, normalized, and combined to form a “superflat” for each spectrograph camera, by stacking the 320 normalized flat-field spectra and performing an iterative least-squares B-spline fit with outlier rejection on this  $320 \times 2048$  oversampled data set to obtain an effectively continuous function. For each fiber, the superflat is resampled at every pixel and divided into the extracted flat-field spectrum to form the “fiber flat.” In this way, flat-field variations between fibers are removed, as are small-scale features in the wavelength dependence of the system response.

For the science exposures, the object and sky fibers are spatially traced, with tweaking from the flat-field trace, and optimally extracted. In the extraction, the Gaussian fiber profile fitting can also be tweaked from the fiber-flat image. Scattered light is removed by a fourth-order Chebyshev polynomial fit. Outlying pixels are rejected and masked. The extracted spectra are then flat-fielded by dividing by the fiber flats, and the wavelength solution is applied, using a vacuum wavelength scale, adjusting slightly to match the known positions of certain sky lines, and correcting to the

heliocentric frame. The wavelength calibration is accurate to  $10 \text{ km s}^{-1}$  or better.

As with the flats, an oversampled “supersky” is constructed from the 32 sky fibers per plate. For each fiber, the supersky is resampled at every pixel in the object spectrum and subtracted.

Telluric absorption in four wavelength regions in the red spectra is removed using spectrophotometric and reddening standard stars: these are used to construct four “supertelluric” spectra using the B-spline fitting procedure.

Next, spectro2d performs a spectrophotometric flux calibration of the individual exposures: the counts in each exposure are placed on the same scale, and the wavelength scale is rebinned to a simple polynomial in  $\log \lambda$ .

If a smear exposure exists, the counts in each exposure are corrected to match a low-order fit to the smear counts; this corrects for light that falls outside the  $3''$  fiber aperture because of seeing, guiding errors, and atmospheric refraction. However, since the smear-exposure procedure was implemented partway through commissioning, many of the plates included in the EDR do not have smear exposures (see Table 4).

Next, the spectra are flux-calibrated by matching the counts in the mean of the high signal-to-noise ratio spectra of the spectrophotometric and reddening standards on the plate and equating this to the synthetic composite F8 subdwarf spectrum from Pickles (1998). This is placed on an absolute scale by matching the synthesized  $r^*$ -fiber magnitudes of these stars to the SDSS photometry.<sup>71</sup> This procedure removes the instrument response as a function of wavelength in each spectrograph camera. The absolute accuracy of the spectrophotometry is not well measured at this point but can probably be trusted to 20% for point sources when smear exposures are available. Note that for extended sources (galaxies), the smear technique effectively corrects the spectrophotometry to a  $5'' \times 8''$  aperture; in the presence of spectral gradients (e.g., nuclear emission lines), the results may be difficult to interpret.

Finally, for each object the individual science frames, both red and blue halves, are stacked and fitted with the iterative B-spline, with inverse-variance weighting. In the process, outliers due to cosmic rays are rejected and masked and errors in the fluxes estimated. The combined, merged spectra are resampled in constant-velocity pixels ( $\log \lambda$ ), with a pixel scale of  $69 \text{ km s}^{-1}$ . If possible, exposures on multiple nights are combined. If a plate is replugged, however, only the exposures with a given plugging are combined.

Mask bits are set for each pixel of the reduced spectra, as described in Table 11. Two masks are given for the final combined spectra: the bits in the OR mask are set if the relevant flag was set in any of the individual exposures entering the combination, while the AND mask indicates that the flag was set in all of the exposures for that pixel. For most purposes, one should examine the OR mask to be safe. Pixels flagged FULLREJECT, NOSKY, NODATA, or BRIGHTSKY, and those with error set identically to zero are likely to be corrupted. Many of these are due to bad columns on the CCDs or cross talk from very bright objects in neighboring fibers (see the flag NEARWHOPPER).

On occasion, the spectra exhibit unphysical wiggles due to problems in the spectrophotometry, especially in the region of the dichroic split ( $\sim 6000 \text{ \AA}$ ); this will be improved in future releases.

#### 4.10.2. Measuring Spectra

The spectro1d pipeline analyzes the combined, merged spectra output by spectro2d and determines object classifications (galaxy, quasar, star, or unknown) and redshifts; it also provides various line measurements and warning flags. The code attempts to measure an emission and absorption redshift independently for every targeted (nonsky) object. That is, to avoid biases, the absorption and emission codes operate independently, and they both operate independently of any target selection information.

The spectro1d pipeline performs a sequence of tasks for each object spectrum on a plate: The spectrum and error array are read in, along with the pixel mask. Pixels with mask bits set to FULLREJECT, NOSKY, NODATA, or BRIGHTSKY are given no weight in the spectro1d routines. The continuum is then fitted with a fifth-order polynomial, with iterative rejection of outliers (e.g., strong lines). The fit continuum is subtracted from the spectrum. The continuum-subtracted spectra are used for cross-correlating with the stellar templates.

##### 4.10.2.1. Emission-Line Redshifts

Emission lines (peaks in the one-dimensional spectrum) are found by carrying out a wavelet transform of the continuum-subtracted spectrum  $f_c(\lambda)$ :

$$w(a, b) = \frac{1}{\sqrt{b}} \int_{-\infty}^{+\infty} f_c(\lambda) \bar{g}\left(\frac{\lambda - a}{b}\right) d\lambda, \quad (24)$$

where  $g(x; a, b)$  is the wavelet (with complex conjugate  $\bar{g}$ ) with translation and scale parameters  $a$  and  $b$ . We apply the *à trous* wavelet (Starck, Siebenmorgen, & Gredel 1997). For fixed wavelet scale  $b$ , the wavelet transform is computed at each pixel center  $a$ ; the scale  $b$  is then increased in geometric steps and the process repeated. Once the full wavelet transform is computed, the code finds peaks above a threshold and eliminates multiple detections (at different  $b$ ) of a given line by searching nearby pixels. The output of this routine is a set of positions of candidate emission lines.

This list of lines with nonzero weights is matched against a list of common galaxy and quasar emission lines, given in Table 30, many of which were measured from the composite quasar spectrum of Vanden Berk et al. (2001; because of velocity shifts of different lines in quasars, the wavelengths listed do not necessarily match their rest-frame values). Each significant peak found by the wavelet routine is assigned a trial line identification from the common list (e.g., Mg II) and an associated trial redshift. The peak is fitted with a Gaussian, and the line center, width, and height above the continuum are stored in the SpecLine class as parameters `wave`, `sigma`, and `height`, respectively. If the code detects close neighboring lines, it fits them with multiple Gaussians. Depending on the trial line identification, the line width it tries to fit is physically constrained. The code then searches for the other expected common emission lines at the appropriate wavelengths for that trial redshift and computes a confidence level (CL) by summing over the weights of the found lines and dividing by the summed weights of the expected lines. The CL is penalized if the dif-

<sup>71</sup> In future data releases, we expect to improve on this technique by carrying out multiple smear exposures of the fundamental SDSS standard stars.

TABLE 30  
REFERENCE WAVELENGTHS, IN VACUUM,  
USED IN `spectroid`

$\lambda_{\text{vac}}$ (Å)	WEIGHT		
	Galaxy	Quasar	SPECIES
Emission Lines			
1033.30 ....	0.0	1.0	O VI
1215.24 ....	0.0	9.0	Ly $\alpha$
1239.42 ....	0.0	3.0	N V
1305.53 ....	0.0	0.0	O I
1335.52 ....	0.0	0.0	C II
1397.61 ....	0.0	0.0	Si IV
1399.8 ....	0.0	1.0	Si IV + O IV
1545.86 ....	0.0	8.0	C IV
1637.85 ....	0.0	0.0	He II
1665.85 ....	0.0	0.0	O III
1857.4 ....	0.0	0.0	Al III
1908.27 ....	0.0	7.0	C III
2326.0 ....	0.0	0.5	C II
2439.5 ....	0.0	0.0	Ne IV
2800.32 ....	1.0	8.0	Mg II
3346.79 ....	0.0	0.0	Ne V
3426.85 ....	0.0	0.0	Ne VI
3728.30 ....	5.0	1.0	O II
3889.0 ....	0.0	0.0	He I
3971.19 ....	0.0	0.0	He
4072.3 ....	0.0	0.0	S II
4102.89 ....	0.5	2.0	H $\delta$
4341.68 ....	1.0	3.0	H $\gamma$
4364.436....	0.0	0.0	O III
4862.68 ....	2.0	4.0	H $\beta$
4932.603....	0.0	0.0	O III
4960.295....	2.0	2.0	O III
5008.240....	3.0	2.0	O III
6302.046....	0.0	0.0	O I
6365.536....	0.0	0.0	O I
6529.03 ....	0.0	0.0	N I
6549.86 ....	3.0	0.0	N II
6564.61 ....	8.0	8.0	H $\alpha$
6585.27 ....	3.0	0.0	N II
6718.29 ....	3.0	0.0	S II
6732.67 ....	3.0	0.0	S II
Absorption Lines			
3934.777...	-1.0	0.0	K
3969.588...	-1.0	0.0	H
4305.61 ....	-1.0	0.0	G
5176.7 ....	-1.0	0.0	Mg
5895.6 ....	-1.0	0.0	Na
Sky Lines			
5578.5 ....	0.0	0.0	Sky
5894.6 ....	0.0	0.0	Sky
6301.7 ....	0.0	0.0	Sky
7246.0 ....	0.0	0.0	Sky

ferent line centers do not quite match. Once all of the trial line identifications and redshifts have been explored, an emission-line redshift is chosen as the one with the highest CL and stored as  $z$  in the `EmissionRedshift` class. The exact expression for the emission-line CL has been tweaked to match our empirical success rate in assigning correct emission-line redshifts, based on manual inspection of a large number of spectra from the EDR (§ 4.10.3).

The `SpecLine` class also gives the errors, continuum, equivalent width,  $\chi^2$ , spectral index, and significance of each line, in parameters listed in Table 13. We caution that the emission-line measurement for H $\alpha$  should only be used if  $\chi^2_{\text{dof}} < 2.5$ . In the `SpecLine` class, “found” lines denote only those lines used to measure the emission-line redshift, while “measured” lines are all lines in the emission-line list (Table 30) measured at the redshifted positions appropriate to the final redshift assigned to the object.

A separate routine searches for high-redshift ( $z > 2.3$ ) quasars by identifying spectra that contain a Ly $\alpha$  forest signature: a broad emission line with more fluctuation on the blue side than on the red side of the line. The routine outputs the wavelength of the Ly $\alpha$  emission line; while this allows a determination of the redshift, it is not a high-precision estimate, because the Ly $\alpha$  line is intrinsically broad and affected by Ly $\alpha$  absorption. The `spectro1d` pipeline stores this as an additional emission-line redshift. This redshift information is stored in the `EmissionRedshift` class (Table 13), which includes an association to an object in the `SpecObj` class. Thus one `SpecObj` can have one or two `EmissionRedshift` objects associated with it.

If the highest CL emission-line redshift uses lines only expected for quasars (e.g., Ly $\alpha$ , C IV, C III]; see Table 30), then the object is provisionally classified as a quasar. If any of the identified lines is broader than 500 km s $^{-1}$  (FWHM), then the object is also provisionally classified as a quasar for the EDR. We expect that this threshold will change in future data releases. These provisional classifications will hold up if the final redshift assigned to the object (see below) agrees with its emission redshift.

#### 4.10.2.2. Cross-Correlation Redshift

The spectra are cross-correlated with stellar, emission-line galaxy, and quasar template spectra to determine a cross-correlation redshift and error. The cross-correlation templates are obtained from SDSS commissioning spectra of high signal-to-noise ratio and comprise roughly one for each stellar spectral type from B to almost L, a nonmagnetic and a magnetic white dwarf, an emission-line galaxy, a composite LRG spectrum, and a composite quasar spectrum (from Vanden Berk et al. 2001). The composites are based on co-additions of  $\sim 2000$  spectra each. The template redshifts are determined by cross-correlation with a large number of stellar spectra from SDSS observations of the M67 star cluster, whose radial velocity is precisely known.

When an object spectrum is cross-correlated with the stellar templates, its found emission lines are masked out, i.e., the redshift is derived from the absorption features. The cross-correlation routine follows the technique of Tonry & Davis (1979): the continuum-subtracted spectrum is Fourier-transformed and convolved with the transform of each template. For each template, the three highest cross-correlation function (CCF) peaks are found, fitted with parabolas, and output with their associated confidence limits. The corresponding redshift errors are given by the widths of the CCF peaks. The cross-correlation CLs are empirically calibrated as a function of peak level based on manual inspection of a large number of spectra from the EDR. The final cross-correlation redshift is then chosen as the one with the highest CL from among all of the templates.

If there are discrepant high-CL cross-correlation peaks, i.e., if the highest peak has CL  $< 0.99$  and the next highest

peak corresponds to a CL that is greater than 70% of the highest peak, then the code extends the cross-correlation analysis for the corresponding templates to lower wave-number and includes the continuum in the analysis, i.e., it chooses the redshift based on which template provides a better match to the continuum shape of the object. These flagged spectra are then manually inspected (see below). The cross-correlation redshift is stored as  $z$  in the `CrossCorrelationRedshift` class.

#### 4.10.2.3. Final Redshifts and Spectrum Classification

The `spectro1d` pipeline assigns a *final* redshift to each object spectrum by choosing the emission or cross-correlation redshift with the highest CL and stores this as  $z$  in the `SpecObj` class. A redshift status bit mask (`zStatus`) and a redshift warning bit mask (`zWarning`) are stored, with bit mask values given in Table 12. The CL is stored in `zConf`. Objects with redshifts determined manually (see below) have CL set to 0.95 (`MANUAL_HIC` set in `zStatus`), or 0.4 or 0.65 (`MANUAL_LOC` set in `zStatus`). Rarely, objects have the entire red or blue half of the spectrum missing; such objects have their CLs reduced by a factor of 2, so they are automatically flagged as having low confidence, and the mask bit `Z_WARNING_NO_BLUE` or `Z_WARNING_NO_RED` is set in `zWarning` as appropriate.

All objects are classified (in `specClass`) as either a quasar, high-redshift quasar, galaxy, star, late-type star, or unknown. If the object has been identified as a quasar by the emission-line routine, and if the emission-line redshift is chosen as the final redshift, then the object retains its quasar classification. Also, if the quasar cross-correlation template provides the final redshift for the object, then the object is classified as a quasar. If the object has a final redshift  $z > 2.3$  (so that  $\text{Ly}\alpha$  is or should be present in the spectrum), and if at least two out of three redshift estimators agree on this (the three estimators being the emission-line,  $\text{Ly}\alpha$ , and cross-correlation redshifts), then it is classified as a high- $z$  quasar. If the object has a redshift  $cz < 450 \text{ km s}^{-1}$ , then it is classified as a star. If the final redshift is obtained from one of the late-type stellar cross-correlation templates, it is classified as a late-type star. If the object has a cross-correlation  $\text{CL} < 0.25$ , it is classified as unknown.

There exist among the spectra a small number of composite objects. Most common are bright stars on top of galaxies, but there are also galaxy-galaxy pairs at distinct redshifts, and at least one galaxy-quasar pair, and one galaxy-star pair. Most of these have the `zWarning` flag set, indicating that more than one redshift was found.

#### 4.10.2.4. Additional Spectral Measures

For all objects, Gaussians are fitted at the positions of all expected emission lines in the reference list (see Table 30), not just the common lines, and their parameters are stored in the class `SpecLineIndex` (Table 13).

For galaxies, we compute in the rest frame the equivalent width, magnitude, and their errors for a number of standard line indices, which are useful for stellar population, age, and metallicity measures. These are stored in `ew`, `ewErr`, `mag`, and `magErr` in the class `SpecLineIndex`. The equivalent width and magnitude are not corrected for velocity dispersion. We only recommend the use of these line indices for stellar population studies for galaxy spectra with good sig-

nal-to-noise ratio: `sn[0]` of class `SpecObj` should be greater than 30.

A complete list of the lines we measure is available on our Web sites. They were gathered from the following sources:

1. Lick line indices (21 absorption-line strengths on the revised Lick/IDS line-strength system; Trager et al. 1998);
2. CN, HK, Ca, G,  $\text{H}\beta$ , Mg, MH, FC, and Na (Brodie & Hanes 1986); and
3. Ca II  $\lambda 8498$ ,  $\lambda 8542$ , and  $\lambda 8662$  and Mg I  $\lambda 8807$  (Díaz, Terlevich, & Terlevich 1989).

We also calculate a measure of the strength of the 4000 Å break as the ratio of the flux from 3751 to 3951 Å to the flux from 4051 to 4251 Å, and the ratio of the strength of the Ca II K and H lines, as integrated from 3921 to 3946 Å and 3956 to 3981 Å, respectively. These are also stored as “lines” in the class `SpecLineIndex`, with the flux ratio stored in the parameter `ew`. This flux ratio is used as a sanity check on the spectra. If the break is in the wrong direction (flux *decreasing* with increasing wavelength), or if the flux ratio is greater than 0.95, the `Z_WARNING_4000BREAK` bit is set in `zWarning`.

Finally, galaxies are classified by a principal component analysis (PCA; Connolly et al. 1995b, as extended by Connolly & Szalay 1999), using cross-correlation with eigentemplates constructed from early SDSS spectroscopic data. The five eigencoefficients and a classification number are stored in `eCoeff` and `eClass`, respectively, in the `SpecObj` class. `eClass` ranges from about  $-0.35$  to  $0.5$  for early- to late-type galaxies.

In future data releases, we expect to implement more detailed stellar spectral classification based on a larger set of stellar templates. The velocity dispersion of each galaxy is currently a placeholder. It will be calculated and stored in a future release.

#### 4.10.2.5. Redshift Warning Flags

The `zWarning` bit mask mentioned above records problems that the `spectro1d` pipeline found with each spectrum. It provides compact information about the spectra for end users, and it is also used to trigger manual inspection of a subset of spectra on every plate. The bit masks are listed in Table 12. Users should particularly heed warnings about parts of the spectrum missing, low signal-to-noise ratio in the spectrum, significant discrepancies between the various measures of the redshift, and especially low confidence in the redshift determination. In addition, redshifts for objects with `zStatus = FAILED` should not be used.

#### 4.10.2.6. Manual Inspection of Spectra

A small percentage of spectra on every plate are inspected manually, and if necessary, the redshift, classification, `zStatus`, and `CL` are corrected. We inspect those spectra that have `zWarning` or `zStatus` indicating that there were multiple high-confidence cross-correlation redshifts, that the redshift was high ( $z > 3.2$  for a quasar or  $z > 0.5$  for a galaxy), that the confidence was low, that signal-to-noise ratio was low in  $r$ , or that the spectrum was not measured. All objects with `zStatus = EMLINE_HIC` or `EMLINE_LOC`, i.e., for which the redshift was determined only by emission lines, are also examined. If, however, the object has a final  $\text{CL} > 0.98$  and

`zStatus` of either `XCORR_EMLINE` or `EMLINE_XCORR`, then despite the above, it is not manually checked. All objects with either `specClass = SPEC_UNKNOWN` or `zStatus = FAILED` are manually inspected.

Roughly 8% of the spectra in the EDR were thus inspected, of which about one-eighth, or 1% overall, had the classification, redshift, `zStatus`, or `CL` manually corrected. Such objects are flagged with `zStatus` changed to `MANUAL_HIC` or `MANUAL_LOC`, depending on whether we had high or low confidence in the classification and redshift from the manual inspection. Tests on the validation plates, described in the next section, indicate that this selection of spectra

for manual inspection successfully finds over 95% of the spectra for which the automated pipeline assigns an incorrect redshift.

#### 4.10.3. Spectroscopic Pipelines Testing and Performance

In order to assess the performance of the spectroscopic pipelines, we determined the redshift and classification by hand of every spectrum on 39 plates from the EDR (a total of roughly 23,000 spectra). Comparing with the outputs of the pipeline, after manual correction following the procedure outlined above, we find that 99.7% of galaxies, 97.9% of quasars, and 99.1% of stars are correctly classified, and the redshifts are correct (to within a pixel or two as determined by eye) for 99.7% of the galaxies, 98.0% of the quasars, and 99.6% of the stars.

We have also compared outputs for plates observed with multiple pluggings and have found that redshifts for galaxies in the main sample are reproducible to on the order of 30 km s<sup>-1</sup>. Finally, we have compared our redshifts with those found in the literature (e.g., the Las Campanas Redshift Survey; Shectman et al. 1996), and we find differences consistent with their and our quoted errors. A similar comparison with the ESO Nearby Abell Cluster Survey

TABLE 31  
MANUAL CLASSIFICATION OF  
NONGALAXY SPECTRA

Identification	Number
Quasar .....	4833
Galaxy with absorption.....	3430
Emission-line galaxy .....	1628
White dwarf .....	734
White dwarf–M dwarf pair.....	61
O star .....	4
B star .....	66
A star .....	957
F star .....	2294
G star.....	1126
K star.....	970
M star .....	1186
L star .....	15
Subdwarf star.....	12
Carbon star.....	23
Cataclysmic variable .....	7
Planetary nebula star .....	2
Star of unknown type.....	82
Unknown type .....	311
Defective spectrum.....	60

(ENACS; P. Katgert 2001, private communication) shows a mean offset of  $\sim 40 \pm 20$  km s<sup>-1</sup> and a dispersion on the order of 90 km s<sup>-1</sup>, a large part of which can be accounted for by the ENACS estimated errors, which were confirmed with repeated measurements. A small number of redshifts from Abell 119 have  $|\Delta_z| > 0.002$ . These have been traced to possible problems with redshifts for several objects from a single ENACS pointing.

Richards et al. (2001) and Schneider et al. (2002) compare SDSS quasar redshifts with those from the NASA/IPAC Extragalactic Database (NED), and they find excellent

TABLE 32  
SUMMARY OF `specObj` OBJECTS INCLUDED IN THE REDSHIFT HISTOGRAMS

Object	Main Galaxy <sup>a</sup>	LRG <sup>b</sup>	Main Quasar <sup>c</sup>	High- <i>z</i> Quasar <sup>d</sup>	Other <sup>e</sup>
Targets <sup>f</sup> .....	35838	9714	7876	1307	6398
<i>z</i> confidence <sup>g</sup> .....	35614	9691	7714	1218	5974
Good <code>zStatus</code> <sup>h</sup> ....	33860	9049	7475	1107	5606
Galaxy <sup>i</sup> .....	32949	8872	2783	227	403
Quasar <sup>j</sup> .....	187	7	3157	263	795
Neither <sup>k</sup> .....	724	170	1535	617	4408

NOTE.—There are a total of 54,008 spectra in the EDR. Columns summarize objects in the specified samples, and rows indicate the number of objects that satisfy the criteria. Total number of spectra classified as galaxy: 39,959. Total number of spectra classified as quasar: 4055.

<sup>a</sup> The main galaxy sample has `primTarget = TARGET_GALAXY` or `TARGET_GALAXY_BIG` or `TARGET_GALAXY_BRIGHT_CORE`.

<sup>b</sup> The luminous red galaxy sample has `primTarget = TARGET_GALAXY_RED` or `TARGET_GALAXY_RED_II`.

<sup>c</sup> The main quasar sample has `primTarget = TARGET_QSO_CAP` or `TARGET_QSO_SKIRT` or `TARGET_QSO_FIRST_CAP` or `TARGET_QSO_FIRST_SKIRT`.

<sup>d</sup> The high-redshift quasar sample has `primTarget = TARGET_QSO_HIZ`.

<sup>e</sup> The “other” sample has none of the above `primTarget` flag bits set.

<sup>f</sup> Number of spectra observed in each sample.

<sup>g</sup> Number of spectra with `zConf ≥ 0.60`.

<sup>h</sup> Number of spectra with a good `zStatus`, not equal to `NOT_MEASURED`, `FAILED`, `INCONSISTENT`, `XCORR_LOC`, `EMLINE_LOC`, or `MANUAL_LOC`.

<sup>i</sup> Number of spectra classified as galaxy (`specClass = 2`).

<sup>j</sup> Number of spectra classified as quasar (`specClass = 3` or `4`).

<sup>k</sup> Number of spectra not classified as either galaxy or quasar.

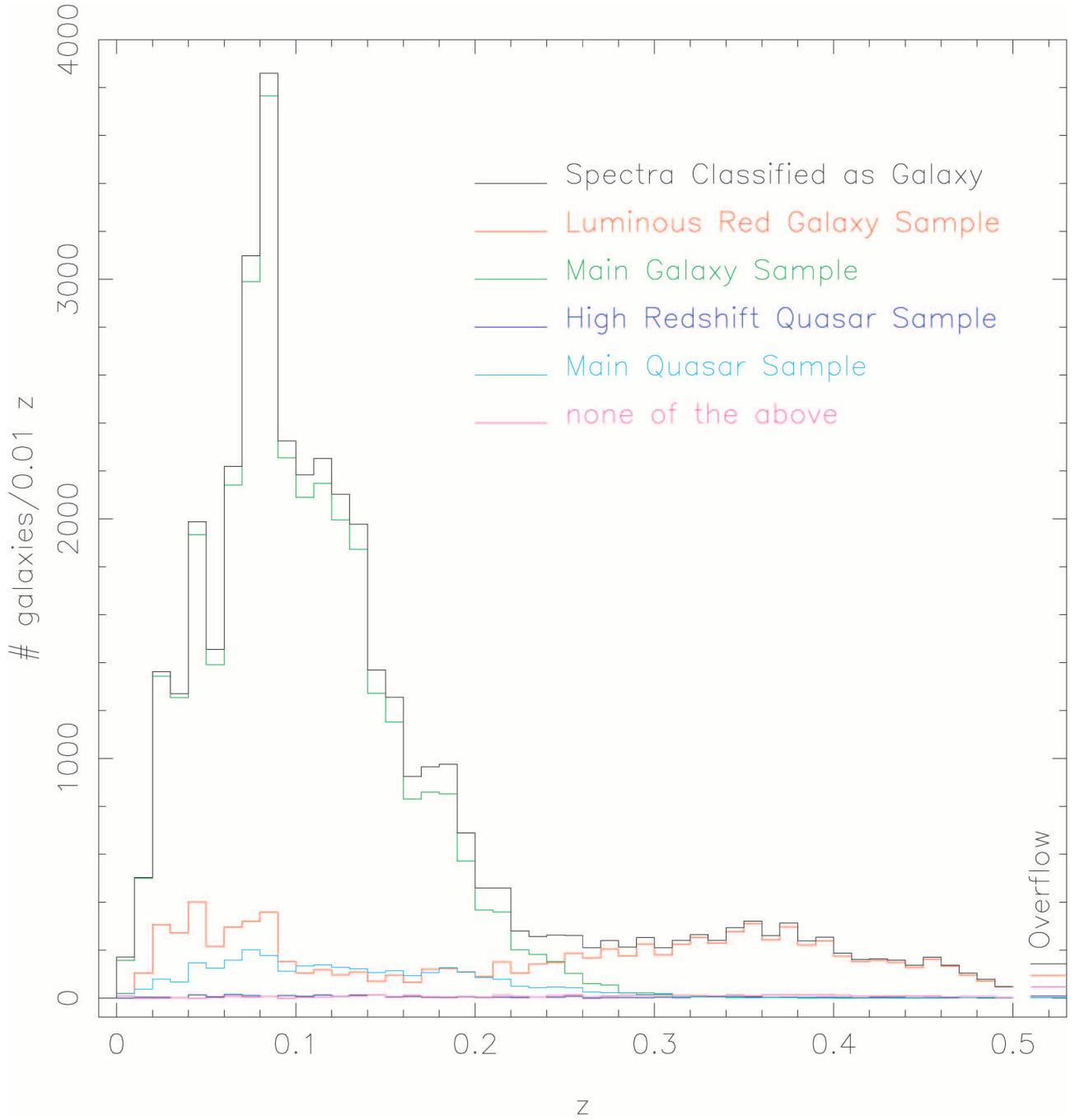


FIG. 14.—Redshift histograms for spectra classified as galaxies

agreement; in three cases, they found that the NED redshifts were wrong, but in no case was the NED redshift correct and the SDSS redshift wrong.

#### 4.11. The Galaxy and Quasar Redshift Samples

We inspected the spectra of all EDR objects that were selected for reasons other than being a galaxy, or that were classified by the spectroscopic pipelines as not being a galaxy. Table 31 summarizes the number of quasars, galaxies, stars of various types, unknown objects, and spectra that are defective for some reason. We have prepared a catalog

of all spectroscopically confirmed quasars in the EDR by visually inspecting all of the spectra. This catalog is presented in Schneider et al. (2002) and is available on our Web sites.

Various spectroscopic samples may be constructed from the EDR. Table 32 summarizes the number of objects with spectra in four broad target samples: “main galaxy,” “luminous red galaxy,” “main quasar,” and “high-redshift quasar.” In the table, we first list the number of targets selected in each class (as well as in all “other” samples), and then the number of spectra with high confidence and good status for the redshift determi-

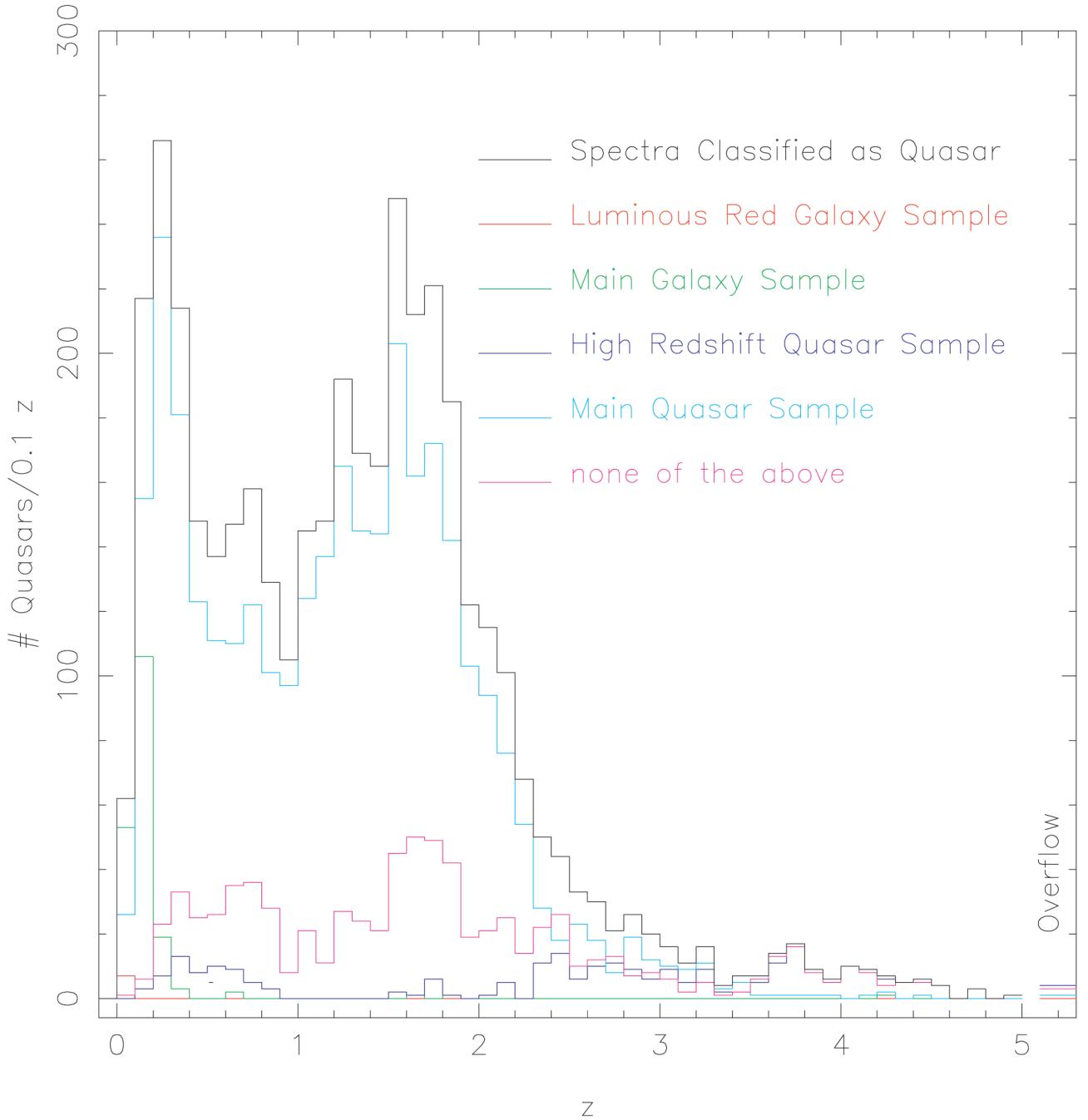


FIG. 15.—Redshift histograms for spectra classified as quasars

nation, respectively. Finally, we list the number of objects classified as galaxy, quasar, and neither, for each of the samples.

Figures 14 and 15 show the redshift distributions for the galaxy and quasar samples, respectively. The contribution from each of the samples is shown, along with the distribution of all redshifts for classified spectra. Figures 16 and 17 are “pie diagrams” for objects from the equatorial data (runs 94/125 and 752/756).

##### 5. BEYOND THE EARLY DATA RELEASE

We encourage the active use of these data from the SDSS EDR. We have worked to make the data products as accu-

rate as possible, but we realize that some calibration or cataloging errors may have crept into our work. Feedback to the help desk set up at STScI will allow us to evaluate the usefulness and correctness of these catalogs. As appropriate, we will update the calibration and contents of this early data release and announce these changes on the users’ group mailing list and the SDSS Web site.

The next release of SDSS data is scheduled for January of 2003. It is scheduled to include at least 1500 deg<sup>2</sup> of imaging data and corresponding spectra.

The Sloan Digital Sky Survey is a joint project of the University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins Uni-

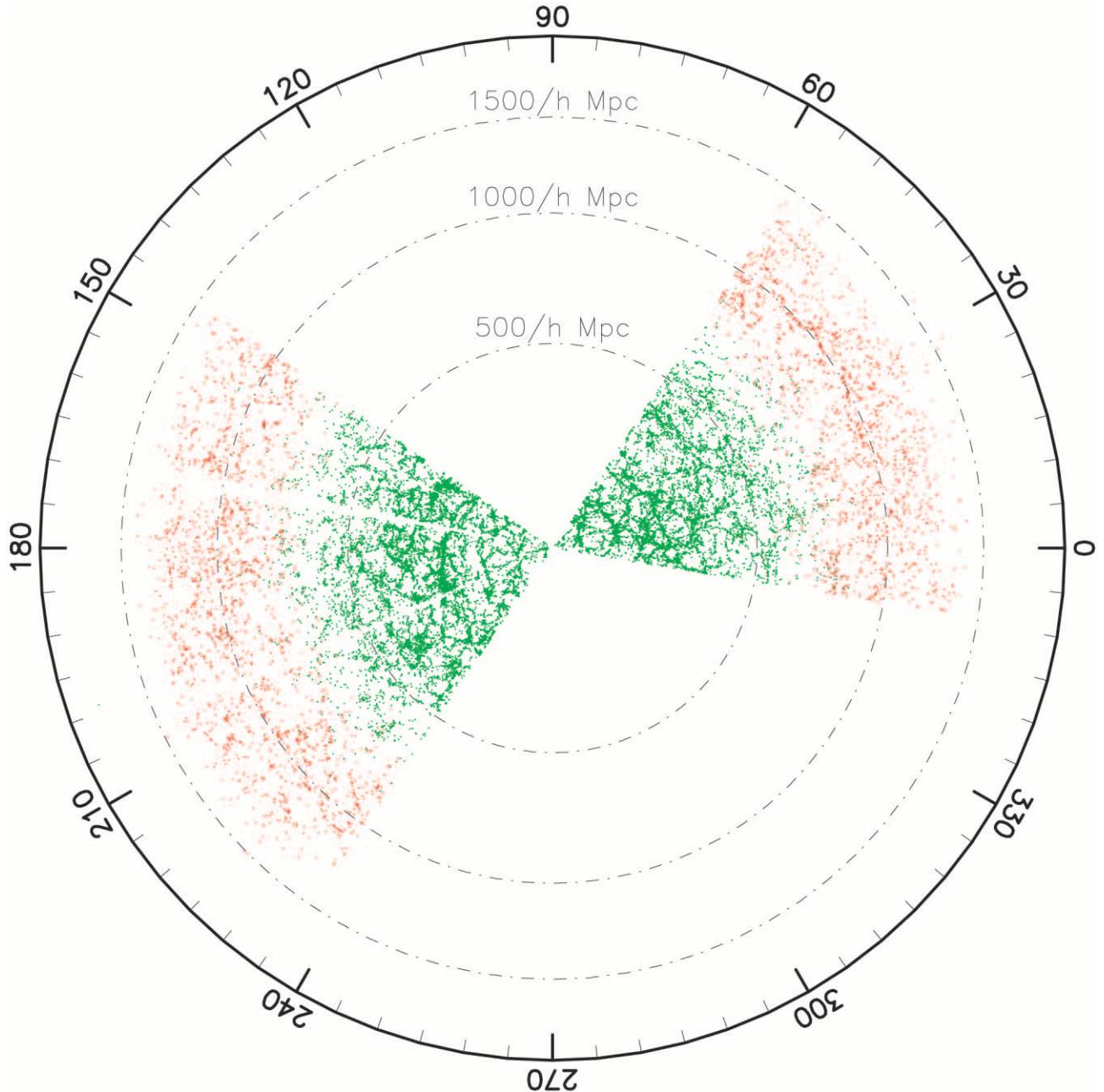


FIG. 16.—Pie diagram of main galaxies (green) and LRGs (red) with spectra in the EDR. The annuli are labeled with the comoving distance, assuming  $\Lambda = 0.7$ ,  $\Omega_M = 0.3$ , and  $H_0 = 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

versity, the Max-Planck-Institut für Astronomie, the Max-Planck-Institut für Astrophysik, New Mexico State University, Princeton University, the US Naval Observatory, and the University of Washington. Apache Point Observatory, site of the SDSS telescopes, is operated by the Astrophysical Research Consortium. Funding for the project has been

provided by the Alfred P. Sloan Foundation, the SDSS member institutions, the National Aeronautics and Space Administration, the National Science Foundation, the US Department of Energy, the Japanese Monbukagakusho, and the Max-Planck-Gesellschaft. The SDSS Web site is <http://www.sdss.org/>.

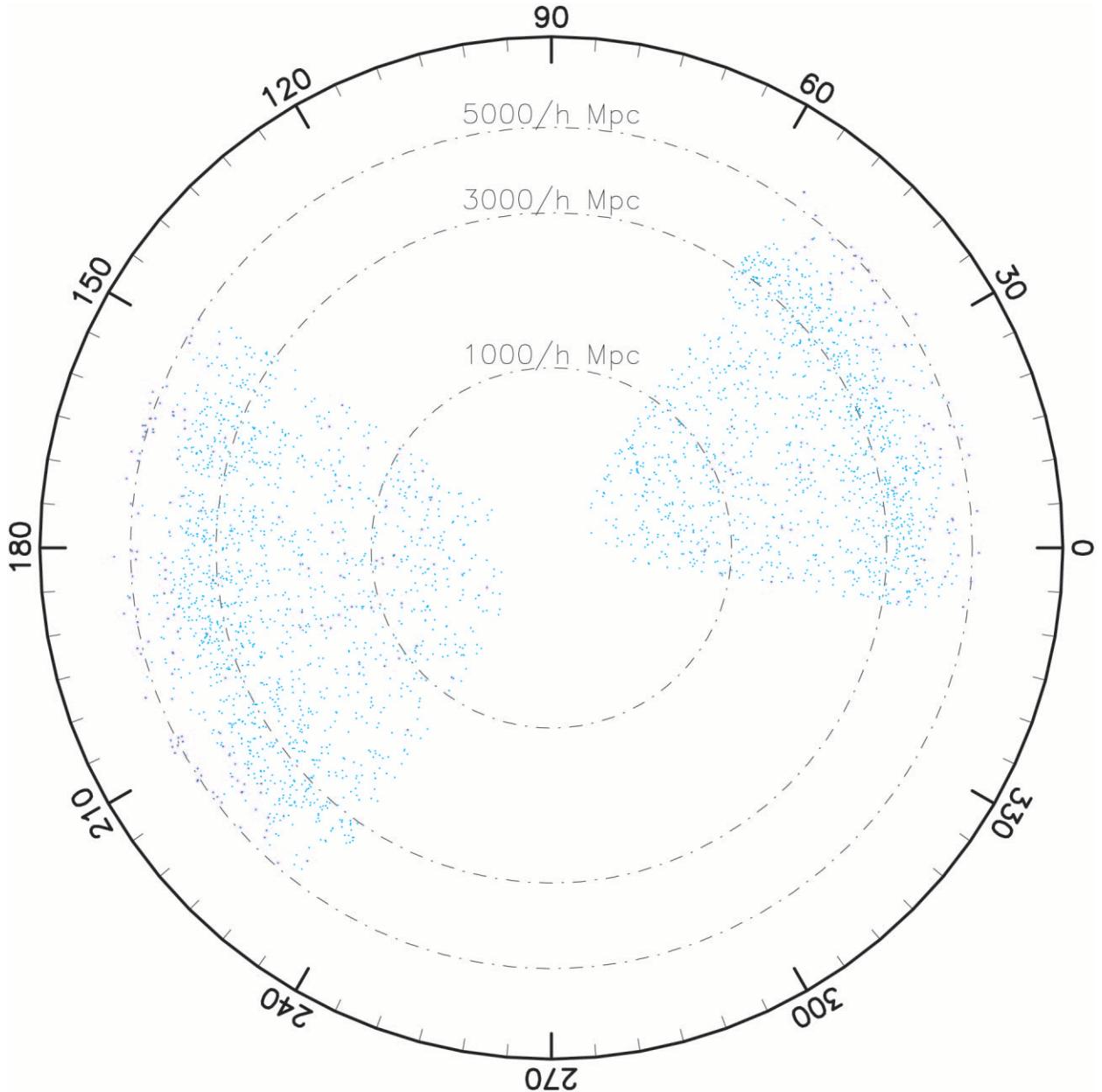


FIG. 17.—Same as Fig. 15, but for quasars: the main sample (green) and the high-redshift sample (blue)

#### REFERENCES

- Anderson, S. F., et al. 2001, AJ, 122, 503  
 Bahcall, J. N., & Soneira, R. M. 1980, ApJS, 44, 73  
 Becker, R. H., White, R. L., & Helfand, D. J. 1995, ApJ, 450, 559  
 Blanton, M. R., et al. 2001, AJ, 121, 2358  
 Brodie, J. P., & Hanes, D. A. 1986, ApJ, 300, 258  
 Castander, F. J., et al. 2001, AJ, 121, 2331  
 Chen, B., et al. 2001, ApJ, 553, 184  
 Connolly, A. J., Csabai, I., Szalay, A. S., Koo, D. C., Kron, R. G., & Munn, J. A. 1995a, AJ, 110, 2655  
 Connolly, A. J., & Szalay, A. S. 1999, AJ, 117, 2052  
 Connolly, A. J., Szalay, A. S., Bershadsky, M. A., Kinney, A. L., & Calzetti, D. 1995b, AJ, 110, 1071  
 Delaney, K. 2000, Inside Microsoft SQL Server 2000 (Redmond, WA: Microsoft Press)  
 Diaz, A. I., Terlevich, E., & Terlevich, R. 1989, MNRAS, 239, 325  
 Eisenstein, D. J., et al. 2001, AJ, 122, 2267  
 Fan, X. 1999, AJ, 117, 2528  
 Fan, X., et al. 2001a, AJ, 122, 2833  
 ———. 2001b, AJ, 121, 31  
 ———. 1999, AJ, 118, 1  
 Finlator, K., et al. 2000, AJ, 120, 2615  
 Fischer, P., et al. 2000, AJ, 120, 1198  
 Fukugita, M., Ichikawa, T., Gunn, J. E., Doi, M., Shimasaku, K., & Schneider, D. P. 1996, AJ, 111, 1748  
 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. 2000, AJ, 120, 963  
 ———. 2001, AJ, 122, 2749  
 Kunszt, P. Ž., Szalay, A. S., Csabai, I., & Thakar, A. R. 2000, in ASP Conf. Ser. 216, Astronomical Data Analysis Software and Systems IX, ed. N. Manset, C. Veillet, & D. Crabtree (San Francisco: ASP), 141  
 Leggett, S. K., et al. 2000, ApJ, 536, L35  
 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  
 Menou, K., et al. 2001, ApJ, 561, 645

- Monet, D. G. 1998, BAAS, 193, No. 120.03
- Newberg, H. J., & Yanny, B. 1997, ApJS, 113, 89
- Objectivity, Inc. 2001, Objectivity Database Reference Manual (Menlo Park, CA: Objectivity, Inc.)
- Petravick, D., et al. 1994, Proc. SPIE, 2198, 935
- Petrosian, V. 1976, ApJ, 209, L1
- Pickles, A. J. 1998, PASP, 110, 863
- Pogson, N. 1856, MNRAS, 17, 12
- Postman, M., & Lauer, T. R. 1995, ApJ, 440, 28
- Richards, G. T., et al. 2001, AJ, 121, 2308  
\_\_\_\_\_. 2002, in preparation
- Rybicki, G. B., & Press, W. H. 1992, ApJ, 398, 169
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
- Schneider, D. P., et al. 2002, AJ, in press
- Scranton, R., et al. 2001, ApJ, submitted (astro-ph/0107416)
- Shectman, S. A., Landy, S. D., Oemler, A., Tucker, D. L., Lin, H., Kirshner, R. P., & Schechter, P. L. 1996, ApJ, 470, 172
- Sheldon, E. S., et al. 2001, ApJ, 554, 881
- Shimasaku, K., et al. 2001, AJ, 122, 1238
- Smith, J. A., et al. 2002, AJ, in press
- Starck, J.-L., Siebenmorgen, R., & Gredel, R. 1997, ApJ, 482, 1011
- Strateva, I., et al. 2001, AJ, 122, 1861
- Szalay, A. S., Gray, J., Kunszt, P., Thakar, A., & Slutz, D. 2001, in Mining the Sky, ed. A. J. Banday, S. Zaroubi, & M. Bartelmann (Berlin: Springer), 99
- Thakar, A. R., Kunszt, P. Z., Szalay, A. S., & Szokoly, G. P. 2000, in ASP Conf. Ser. 216, Astronomical Data Analysis Software and Systems IX, ed. N. Manset, C. Veillet, & D. Crabtree (San Francisco: ASP), 231
- Tonry, J., & Davis, M. 1979, AJ, 84, 1511
- Trager, S. C., Worthey, G., Faber, S. M., Burstein, D., & González, J. J. 1998, ApJS, 116, 1
- Vanden Berk, D. E., et al. 2001, AJ, 122, 549
- Voges, W., et al. 1999, A&A, 349, 389
- Waddell, P., Mannery, E. J., Gunn, J. E., & Kent, S. M. 1998, Proc. SPIE, 3352, 742
- Yanny, B., et al. 2000, ApJ, 540, 825
- Yasuda, N., et al. 2001, AJ, 122, 1104
- York, D. G., et al. 2000, AJ, 120, 1579
- Zacharias, N., et al. 2000, AJ, 120, 2131
- Zehavi, I., et al. 2001, ApJ, submitted (astro-ph/0106476)