The Astrophysical Journal Supplement Series, 182:543–558, 2009 June doi:10.1088/0067-0049/182/2/543
© 2009. The American Astronomical Society. All rights reserved. Printed in the U.S.A.

THE SEVENTH DATA RELEASE OF THE SLOAN DIGITAL SKY SURVEY

Kevork N. Abazajian[1], Jennifer K. Adelman-McCarthy[2], Marcel A. Agüeros[3,102], Sahar S. Allam[2,4], Carlos Allende Prieto[5], Deokkeun An[6,7], Kurt S. J. Anderson[8,9], Scott F. Anderson[10], James Annis[2], Neta A. Bahcall[11], C. A. L. Bailer-Jones[12], J. C. Barentine[13], Bruce A. Bassett[14,15], Andrew C. Becker[10], Timothy C. Beers[16], Eric F. Bell[12], Vasily Belokurov[17], Andreas A. Berlind[18], Eileen F. Berman[2], Mariangela Bernardi[19], Steven J. Bickerton[11], Dmitry Bizyaev[8], John P. Blakeslee[20], Michael R. Blanton[21], John J. Bochanski[10,22], William N. Boroski[2], Howard J. Brewington[8], Jarle Brinchmann[23,24], J. Brinkmann[8], Robert J. Brunner[25], Tamás Budavári[26], Larry N. Carey[10], Samuel Carliles[26], Michael A. Carr[11], Francisco J. Castander[27], David Cinabro[28], A. J. Connolly[10], István Csabai[29], Carlos E. Cunha[30], Paul C. Czarapata[2], James R. A. Davenport[31], Ernst de Haas[32], Ben Dilday[33,34,35], Mamoru Doi[36,37], Daniel J. Eisenstein[38], Michael L. Evans[10], N. W. Evans[17], Xiaohui Fan[38], Scott D. Friedman[39], Joshua A. Frieman[2,34,40], Masataka Fukugita[41], Boris T. Gänsicke[42], Evalyn Gates[34], Bruce Gillespie[26], G. Gilmore[17], Belinda Gonzalez[2], Carlos F. Gonzalez[2], Eva K. Grebel[43], James E. Gunn[11], Zsuzsanna Györ[29], Patrick B. Hall[44], Paul Harding[45], Frederick H. Harris[46], Michael Harvanek[47], Suzanne L. Hawley[10], Jeffrey J. E. Hayes[48], Timothy M. Heckman[26], John S. Hendry[2], Gregory S. Hennessy[49], Robert B. Hindsley[50], J. Hoblitt[51], Craig J. Hogan[2], David W. Hogg[21], Jon A. Holtzman[9], Joseph B. Hyde[19], Shin-ichi Ichikawa[52], Takashi Ichikawa[53], Myungshin Im[54], Željko Ivezić[10], Sebastian Jester[12], Linhua Jiang[38], Jennifer A. Johnson[6], Anders M. Jorgensen[55], Mario Jurić[56], Stephen M. Kent[2], R. Kessler[34], S. J. Kleinman[57], G. R. Knapp[11], Kohki Konishi[41,58], Richard G. Kron[2,40], Jurek Krzesinski[8,59], Nikolay Kuropatkin[2], Hubert Lampeitl[60], Svetlana Lebedeva[2], Myung Gyoon Lee[54], Young Sun Lee[16], R. French Leger[10], Sébastien Lépine[61], Nolan Li[26], Marcos Lima[19,33,34], Huan Lin[2], Daniel C. Long[8], Craig P. Loomis[11], Jon Loveday[62], Robert H. Lupton[11], Eugene Magnier[51], Olena Malanushenko[8], Viktor Malanushenko[8], Rachel Mandelbaum[56,103], Bruce Margon[63], John P. Marriner[2], David Martínez-Delgado[64], Takahiko Matsubara[65], Peregrine M. McGehee[7], Timothy A. McKay[30], Avery Meiksin[66], Heather L. Morrison[45], Fergal Mullally[11], Jeffrey A. Munn[46], Tara Murphy[66,67], Thomas Nash[2], Ada Nebot[68], Eric H. Neilsen, Jr.[2], Heidi Jo Newberg[69], Peter R. Newman[8,70], Robert C. Nichol[60], Tom Nicinski[2,71], Maria Nieto-Santisteban[26], Atsuko Nitta[57], Sadanori Okamura[72], Daniel J. Oravetz[8], Jeremiah P. Ostriker[11], Russell Owen[10], Nikhil Padmanabhan[73,103], Kaike Pan[8], Changbom Park[74], George Pauls[11], John Peoples Jr.[2], Will J. Percival[60], Jeffrey R. Pier[46], Adrian C. Pope[51,75], Dimitri Pourbaix[11,76], Paul A. Price[51], Norbert Purger[29], Thomas Quinn[10], M. Jordan Raddick[26], Paola Re Fiorentin[12,77], Gordon T. Richards[78], Michael W. Richmond[79], Adam G. Riess[26], Hans-Walter Rix[12], Constance M. Rockosi[80], Masao Sako[19,81], David J. Schlegel[73], Donald P. Schneider[82], Ralf-Dieter Scholz[68], Matthias R. Schreiber[83], Axel D. Schwope[68], Uroš Seljak[73,84,85], Branimir Sesar[10], Erin Sheldon[21,86], Kazu Shimasaku[72], Valena C. Sibley[2], A. E. Simmons[8], Thirupathi Sivarani[16,87], J. Allyn Smith[88], Martin C. Smith[17], Vernesa Smolčić[89], Stephanie A. Snedden[8], Albert Stebbins[2], Matthias Steinmetz[68], Chris Stoughton[2], Michael A. Strauss[11], Mark SubbaRao[40,90], Yasushi Suto[58], Alexander S. Szalay[26], István Szapudi[51], Paula Szkody[10], Masayuki Tanaka[91], Max Tegmark[92], Luis F. A. Teodoro[93], Aniruddha R. Thakar[26], Christy A. Tremonti[12], Douglas L. Tucker[2], Alan Uomoto[94], Daniel E. Vanden Berk[82,95], Jan Vandenberg[26], S. Vidrih[43], Michael S. Vogeley[78], Wolfgang Voges[96], Nicole P. Vogt[9], Yogesh Wadadekar[11,97], Shannon Watters[8,98], David H. Weinberg[6], Andrew A. West[22], Simon D. M. White[99], Brian C. Wilhite[100], Alainna C. Wonders[26], Brian Yanny[2], D. R. Yocum[2], Donald G. York[40,101], Idit Zehavi[45], Stefano Zibetti[12], and Daniel B. Zucker[17]

[1] Department of Physics, University of Maryland, College Park, MD 20742, USA
[2] Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, IL 60510, USA
[3] Columbia Astrophysics Laboratory, 550 West 120th Street, New York, NY 10027, USA
[4] Department of Physics and Astronomy, University of Wyoming, Laramie, WY 82071, USA
[5] Mullard Space Science Laboratory, University College London, Holmbury St Mary, Surrey, RH5 6NT, UK
[6] Department of Astronomy, Ohio State University, 140 West 18th Avenue, Columbus, OH 43210, USA
[7] IPAC, MS 220-6, California Institute of Technology, Pasadena, CA 91125, USA
[8] Apache Point Observatory, P.O. Box 59, Sunspot, NM 88349, USA
[9] Department of Astronomy, MSC 4500, New Mexico State University, P.O. Box 30001, Las Cruces, NM 88003, USA
[10] Department of Astronomy, University of Washington, Box 351580, Seattle, WA 98195, USA
[11] Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA
[12] Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany
[13] McDonald Observatory and Department of Astronomy, The University of Texas, 1 University Station, C1400, Austin, TX 78712-0259, USA
[14] South African Astronomical Observatory, Observatory, Cape Town, South Africa
[15] University of Cape Town, Rondebosch, Cape Town, South Africa
[16] Department of Physics & Astrophysics, CSCE: Center for the Study of Cosmic Evolution, and JINA: Joint Institute for Nuclear Astrophysics, Michigan State University, E. Lansing, MI 48824, USA
[17] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge, CB3 0HA, UK
[18] Department of Physics and Astronomy, Vanderbilt University, Nashville, TN 37235, USA
[19] Department of Physics and Astronomy, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, USA
[20] Herzberg Institute of Astrophysics, National Research Council of Canada, 5071 West Saanich Road, Victoria, B. C., V9E 2E7, Canada
[21] Center for Cosmology and Particle Physics, Department of Physics, New York University, 4 Washington Place, New York, NY 10003, USA
[22] MIT Kavli Institute for Astrophysics and Space Research, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
[23] Leiden Observatory, Leiden University, P.O. Box 9513, 2300 RA Leiden, The Netherlands

[24] Centro de Astrofísica da Universidade do Porto, Rua das Estrelas, 4150-762 Porto, Portugal
[25] Department of Astronomy, University of Illinois, 1002 West Green Street, Urbana, IL 61801, USA
[26] Center for Astrophysical Sciences, Department of Physics and Astronomy, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218, USA
[27] Institut de Ciències de l'Espai (IEEC/CSIC), Campus UAB, E-08193 Bellaterra, Barcelona, Spain
[28] Department of Physics and Astronomy, Wayne State University, Detroit, MI 48202, USA
[29] Department of Physics of Complex Systems, Eötvös Loránd University, Pf. 32, H-1518 Budapest, Hungary
[30] Departments of Physics and Astronomy, University of Michigan, 450 Church Street, Ann Arbor, MI 48109, USA
[31] Department of Astronomy, San Diego State University, PA 221, 5500 Campanile Drive, San Diego, CA 92182-1221, USA
[32] Joseph Henry Laboratories, Princeton University, Princeton, NJ 08544, USA
[33] Department of Physics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA
[34] Kavli Institute for Cosmological Physics, The University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA
[35] Department of Physics and Astronomy Rutgers, The State University of New Jersey 136 Frelinghuysen Road Piscataway, NJ 08854-8019, USA
[36] Institute of Astronomy, Graduate School of Science, The University of Tokyo, 2-21-1 Osawa, Mitaka, 181-0015, Japan
[37] Institute for the Physics and Mathematics of the Universe, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, 277-8568, Japan
[38] Steward Observatory, 933 North Cherry Avenue, Tucson, AZ 85721, USA
[39] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[40] Department of Astronomy and Astrophysics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA
[41] Institute for Cosmic Ray Research, The University of Tokyo, 5-1-5 Kashiwa, Kashiwa City, Chiba 277-8582, Japan
[42] Department of Physics, University of Warwick, Coventry, CV4 7AL, UK
[43] Astronomisches Rechen-Institut, Zentrum für Astronomie, University of Heidelberg, Mönchhofstrasse 12-14, D-69120 Heidelberg, Germany
[44] Department of Physics & Astronomy, York University, 4700 Keele St., Toronto, ON, M3J 1P3, Canada
[45] Department of Astronomy, Case Western Reserve University, Cleveland, OH 44106, USA
[46] US Naval Observatory, Flagstaff Station, 10391 W. Naval Observatory Road, Flagstaff, AZ 86001-8521, USA
[47] Lowell Observatory, 1400 W Mars Hill Rd, Flagstaff AZ 86001, USA
[48] Heliophysics Division, Science Mission Directorate, NASA Headquarters, 300 E Street SW, Washington, DC 20546-0001, USA
[49] US Naval Observatory, 3540 Massachusetts Avenue NW, Washington, DC 20392, USA
[50] Code 7215, Remote Sensing Division, Naval Research Laboratory, 4555 Overlook Avenue SW, Washington, DC 20392, USA
[51] Institute for Astronomy, 2680 Woodlawn Road, Honolulu, HI 96822, USA
[52] National Astronomical Observatory, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
[53] Astronomical Institute, Tohoku University, Aoba, Sendai 980-8578, Japan
[54] Department of Physics & Astronomy, Seoul National University, Shillim-dong, San 56-1, Kwanak-gu, Seoul 151-742, Korea
[55] Electrical Engineering Department, New Mexico Institute of Mining and Technology, 801 Leroy Place, Socorro, NM 87801, USA
[56] Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, USA
[57] Gemini Observatory, 670 N. A'ohoku Place, Hilo, HI 96720, USA
[58] Department of Physics and Research Center for the Early Universe, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan
[59] Obserwatorium Astronomiczne na Suhorze, Akademia Pedogogiczna w Krakowie, ulica Podchorażych 2, PL-30-084 Kraców, Poland
[60] Institute of Cosmology and Gravitation (ICG), Mercantile House, Hampshire Terrace, University of Portsmouth, Portsmouth, PO1 2EG, UK
[61] Department of Astrophysics, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA
[62] Astronomy Centre, University of Sussex, Falmer, Brighton, BN1 9QH, UK
[63] Department of Astronomy & Astrophysics, University of California, Santa Cruz, CA 95064, USA
[64] Instituto de Astrofísica de Canarias, E38205 La Laguna, Tenerife, Spain
[65] Department of Physics and Astrophysics, Nagoya University, Chikusa, Nagoya 464-8602, Japan
[66] SUPA, Institute for Astronomy, Royal Observatory, University of Edinburgh, Blackford Hill, Edinburgh, EH9 3HJ, UK
[67] Sydney Institute of Astronomy, The University of Sydney, NSW 2006, Australia
[68] Astrophysical Institute Potsdam, An der Sternwarte 16, 14482 Potsdam, Germany
[69] Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, 110 Eighth Street, Troy, NY 12180, USA
[70] 322 Fulham Palace Road, London, SW6 6HS, UK
[71] CMC Electronics Aurora, 84 N. Dugan Rd. Sugar Grove, IL 60554, USA
[72] Department of Astronomy and Research Center for the Early Universe, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan
[73] Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA
[74] Korea Institute for Advanced Study, 87 Hoegiro, Dongdaemun-Gu, Seoul 130-722, Korea
[75] Los Alamos National Laboratory, P.O. Box 1663, Los Alamos, NM 87545, USA
[76] FNRS Institut d'Astronomie et d'Astrophysique, Université Libre de Bruxelles, CP 226, Boulevard du Triomphe, B-1050 Bruxelles, Belgium
[77] Department of Physics, University of Ljubljana, Jadranska 19, 1000 Ljubljana, Slovenia
[78] Department of Physics, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104, USA
[79] Department of Physics, Rochester Institute of Technology, 84 Lomb Memorial Drive, Rochester, NY 14623-5603, USA
[80] UCO/Lick Observatory, University of California, Santa Cruz, CA 95064, USA
[81] Kavli Institute for Particle Astrophysics & Cosmology, Stanford University, P.O. Box 20450, MS29, Stanford, CA 94309, USA
[82] Department of Astronomy and Astrophysics, 525 Davey Laboratory, Pennsylvania State University, University Park, PA 16802, USA
[83] Universidad de Valparaiso, Departamento de Fisica y Astronomia, Valparaiso, Chile
[84] Physics Department, University of California, Berkeley, CA 94720, USA
[85] Institute for Theoretical Physics, University of Zurich, Zurich 8057, Switzerland
[86] Bldg 510 Brookhaven National Laboratory Upton, NY 11973, USA
[87] Department of Astronomy, University of Florida, Bryant Space Science Center, Gainesville, FL 32611-2055, USA
[88] Department of Physics and Astronomy, Austin Peay State University, P.O. Box 4608, Clarksville, TN 37040, USA
[89] California Institute of Technology, 1200 East California Blvd, Pasadena, CA 91125, USA
[90] Adler Planetarium and Astronomy Museum, 1300 Lake Shore Drive, Chicago, IL 60605, USA
[91] European Southern Observatory, Karl-Schwarzschild-Str. 2, D-85748 Garching bei München, Germany
[92] Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[93] Astronomy and Astrophysics Group, Department of Physics and Astronomy, Kelvin Building, University of Glasgow, Glasgow, G12 8QQ, Scotland, UK
[94] Observatories of the Carnegie Institution of Washington, 813 Santa Barbara Street, Pasadena, CA 91101, USA
[95] Department of Physics, Saint Vincent College, 300 Fraser Purchase Road, Latrobe, PA 15650, USA
[96] Max-Planck-Institut für extraterrestrische Physik, Giessenbachstrasse 1, D-85741 Garching, Germany

[97] National Centre for Radio Astrophysics, Tata Institute of Fundamental Research, Post Bag 3, Ganeshkhind, Pune 411007, India
[98] Advanced Technology and Research Center, Institute for Astronomy, 34 Ohia Ku St., Pukalani, HI 96768, USA
[99] Max-Planck-Institut für Astrophysik, Postfach 1, D-85748 Garching, Germany
[100] Department of Physics, Elmhurst College, 190 Prospect Ave., Elmhurst, IL 60126, USA
[101] Enrico Fermi Institute, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA
Received 2008 December 2; accepted 2009 March 19; published 2009 May 18

ABSTRACT

This paper describes the Seventh Data Release of the Sloan Digital Sky Survey (SDSS), marking the completion of the original goals of the SDSS and the end of the phase known as SDSS-II. It includes 11,663 deg^2 of imaging data, with most of the ∼2000 deg^2 increment over the previous data release lying in regions of low Galactic latitude. The catalog contains five-band photometry for 357 million distinct objects. The survey also includes repeat photometry on a 120° long, 2.°5 wide stripe along the celestial equator in the Southern Galactic Cap, with some regions covered by as many as 90 individual imaging runs. We include a co-addition of the best of these data, going roughly 2 mag fainter than the main survey over 250 deg^2. The survey has completed spectroscopy over 9380 deg^2; the spectroscopy is now complete over a large contiguous area of the Northern Galactic Cap, closing the gap that was present in previous data releases. There are over 1.6 million spectra in total, including 930,000 galaxies, 120,000 quasars, and 460,000 stars. The data release includes improved stellar photometry at low Galactic latitude. The astrometry has all been recalibrated with the second version of the USNO CCD Astrograph Catalog, reducing the rms statistical errors at the bright end to 45 milliarcseconds per coordinate. We further quantify a systematic error in bright galaxy photometry due to poor sky determination; this problem is less severe than previously reported for the majority of galaxies. Finally, we describe a series of improvements to the spectroscopic reductions, including better flat fielding and improved wavelength calibration at the blue end, better processing of objects with extremely strong narrow emission lines, and an improved determination of stellar metallicities.

Key words: atlases – catalogs – surveys

Online-only material: color figures

1. OVERVIEW OF THE SLOAN DIGITAL SKY SURVEY

The Sloan Digital Sky Survey (SDSS; York et al. 2000) saw first light a decade ago, with the goals of obtaining CCD imaging in five broad bands over 10,000 deg^2 of high-latitude sky, and spectroscopy of a million galaxies and 100,000 quasars over this same region. With this, its seventh public data release, these goals have been realized. The survey facilities have also been used to carry out a comprehensive imaging and spectroscopic survey to explore the structure, composition, and kinematics of the Milky Way Galaxy (Sloan Extension for Galactic Understanding and Exploration (SEGUE); Yanny et al. 2009), and a repeat imaging survey that has discovered more than 500 spectroscopically confirmed Type Ia supernovae with superb light curves (Frieman et al. 2008; Holtzman et al. 2008).

The SDSS uses a dedicated wide-field 2.5 m telescope (Gunn et al. 2006) located at Apache Point Observatory (APO) near Sacramento Peak in Southern New Mexico. The telescope uses two instruments. The first is a wide-field imager (Gunn et al. 1998) with 24 2048 × 2048 CCDs on the focal plane with 0.″396 pixels that covers the sky in drift-scan mode in five filters in the order *riuzg* (Fukugita et al. 1996). The imaging is done with the telescope tracking great circles at the sidereal rate; the effective exposure time per filter is 54.1 s, and 18.75 deg^2 are imaged per hour in each of the five filters. The images are mostly taken under good seeing conditions (the median is about 1.″4 in *r*) on moonless photometric nights (Hogg et al. 2001); the exceptions are a series of repeat scans of the celestial equator in the Fall for a supernova search (Frieman et al. 2008), as is described in more detail in Section 3.2. The 95% completeness limits of the images are u, g, r, i, z = 22.0, 22.2, 22.2, 21.3, 20.5, respectively (Abazajian et al. 2004), although these values depend as expected on seeing and sky brightness. The images are processed through a series of pipelines that determine an astrometric calibration (Pier et al. 2003) and detect and measure the brightnesses, positions, and shapes of objects (Lupton et al. 2001; Stoughton et al. 2002). The astrometry is good to 45 milliarcseconds (mas) rms per coordinate at the bright end, as described in more detail in Section 4.4. The photometry is calibrated to an AB system (Oke & Gunn 1983), and the zero points of the system are known to 1%–2% (Abazajian et al. 2004). The photometric calibration is done in two ways, by tying to photometric standard stars (Smith et al. 2002) measured by a separate 0.5 m telescope on the site (Tucker et al. 2006; Ivezić et al. 2004), and by using the overlap between adjacent imaging runs to tie the photometry of all the imaging observations together, in a process called ubercalibration (Padmanabhan et al. 2008). Results of both processes are made available; with this data release, the ubercalibration results, which are uncertain at the ∼1% level in $griz$ and 2% in u, are now the default photometry made available in the data release described in this paper.

The photometric catalogs of detected objects are used to identify objects for spectroscopy with the second of the instruments on the telescope: a 640-fiber-fed pair of multiobject double spectrographs, giving coverage from 3800 Å to 9200 Å at a resolution of $\lambda/\Delta\lambda \simeq 2000$. The objects chosen for spectroscopic follow-up are selected based on photometry corrected for Galactic extinction following Schlegel et al. (1998; hereafter SFD) and include:

1. A sample of galaxies complete to a Petrosian (1976) magnitude limit of $r = 17.77$ (Strauss et al. 2002).
2. Two deeper samples of luminous red ellipticals selected in color–magnitude space to $r = 19.2$ and $r = 19.5$,

[102] NSF Astronomy and Astrophysics Postdoctoral Fellow.
[103] Hubble Fellow.

respectively, which produce an approximately volume-limited sample to $z = 0.38$, and a flux-limited sample extending to $z = 0.55$, respectively (Eisenstein et al. 2001).
3. Flux-limited samples of quasar candidates, selected by their nonstellar colors or FIRST (Becker et al. 1995) radio emission to $i = 19.1$ in regions of color space characteristic of $z < 3$ quasars, and to $i = 20.2$ for quasars with $3 < z < 5.5$ (Richards et al. 2002).
4. A variety of ancillary samples, including optical counterparts to *ROSAT*-detected X-ray sources (Anderson et al. 2007).
5. Stars for spectrophotometric calibration and telluric absorption correction, as well as regions of blank sky for accurate sky subtraction.
6. A variety of categories of stellar targets with a series of color and magnitude cuts for measurements of radial velocity, metallicity, surface temperature, and Galactic structure as part of SEGUE (Yanny et al. 2009).

These targets are arranged on *tiles* of radius 1.°49, with centers chosen to maximize the number of targeted objects (Blanton et al. 2003). Each tile contains 640 objects, and forms the template for an aluminum spectroscopic plate, in which holes are drilled to hold optical fibers that feed the spectrographs. Spectroscopic exposures are 15 minutes long, and three or more are taken for a given plate to reach predefined requirements of signal-to-noise ratio (S/N), namely $(S/N)^2 > 15$ per 1.5 Å pixel for stellar objects of fiber magnitude $g = 20.2$, $r = 20.25$, and $i = 19.9$. For the SEGUE faint plates, the exposures are considerably deeper, and typically consist of eight 15 minute exposures, giving $(S/N)^2 \sim 100$ at the same depth (Yanny et al. 2009).

Spectra are extracted and calibrated in wavelength and flux. The typical S/N of a galaxy near the main sample flux limit is 10 per pixel. The broadband spectrophotometric calibration is accurate to 4% rms for point sources (Adelman-McCarthy et al. 2008), and the wavelength calibration is good to 2 km s^{-1}. The spectra are classified and redshifts determined using a pair of pipelines (Stoughton et al. 2002; Subbarao et al. 2002), which give consistent results 98% of the time; the discrepant objects tend to be of very low S/N, or very unusual objects, such as extreme broad absorption line quasars, superposed sources, and so on. The vast majority of the spectra of galaxies and quasars yield reliable redshifts; the failure rate is of order 1% for galaxies and slightly larger for quasars. The stellar targets are further processed by a separate pipeline (Lee et al. 2008a, 2008b; Allende Prieto et al. 2008a) which determines surface temperatures, metallicities, and gravities.

The resulting catalogs are stored and distributed via a database accessible on the web (the Catalog Archive Server (CAS);[104] Thakar et al. 2008), and the images and flat files are available in bulk through the Data Archive Server (DAS).[105]

The SDSS saw first light in 1998 May and started routine operations in 2000 April. It was originally funded for five years of operations, but had not completed its core goals of imaging and spectroscopy of a large contiguous area of the Northern Galactic Cap by 2005. The survey was extended for an additional three years, with the additional goals of the SEGUE and the supernova surveys mentioned above. The extended program is known as SDSS-II, and the component of SDSS-II that represents the completion of SDSS-I is known as the Legacy Survey. SDSS-II observations were completed in 2008 July.

The SDSS data have been made public in a series of yearly data releases (Stoughton et al. 2002; Abazajian et al. 2003, 2004, 2005; Adelman-McCarthy et al. 2006, 2007, 2008; hereafter the EDR, DR1, DR2, DR3, DR4, DR5, and DR6 papers, respectively). The most recent of these papers described the Sixth Data Release (DR6), which included data taken through 2006 July. The present paper describes the Seventh Data Release (DR7), including data taken through the end of SDSS-II in 2008 July, and thus represents two additional years of data. The data releases are cumulative; DR7 includes all data included in the previous releases as well. In Section 2, we describe the footprint of this survey; most importantly, we have completed our goals of

1. contiguous imaging and spectroscopy over 7500 deg^2 of the Northern Galactic Cap (the Legacy Survey);
2. imaging and spectroscopy of stellar sources over an additional 3500 deg^2 at lower Galactic latitudes to study the structure of the Milky Way; and
3. repeat imaging of >250 deg^2 on the celestial equator in the Fall months to discover Type Ia supernovae with $0.1 < z < 0.4$.

In Section 3, we describe the repeat scans on the celestial equator, including a co-addition of the images to reach about 2 mag deeper than the main survey. In Section 4, we present improvements in the processing of the imaging data, including improved stellar photometry at low Galactic latitudes, an astrometric recalibration, and improvements in our photometric redshift algorithms for galaxies. The DR6 paper described a problem with the photometry of bright galaxies; we explore this further in Section 5. In Section 6, we discuss improvements in the spectroscopic processing of the data. The DR6 paper described improvements in the wavelength and spectrophotometric calibration; we have implemented further refinements which are important in the determination of accurate stellar parameters from the spectra.

We conclude in Section 7 with a discussion of the future of the SDSS project.

2. SURVEY FOOTPRINT

Table 1 summarizes the contents of DR7, giving the imaging and spectroscopic sky coverage and number of objects. The imaging footprint has increased by roughly 22% since DR6 (most of it outside the contiguous area of the North Galactic Cap), and the number of spectra has increased by 29%.

The imaging for the Legacy Survey was substantially complete with DR6. In DR7, we include imaging of a few small gaps that were missed in the contiguous region of the North Galactic Cap, and repeat observations of a few regions of the sky which had particularly poor seeing in previous data releases. The total footprint has increased by less than 10 deg^2 in total. The Legacy imaging footprint is visible as the large contiguous gray area on the left side of the upper panel of Figure 1, together with the three gray stripes visible on the right side. The principal augmentation of the imaging data in DR7 is the stripes which are part of the SEGUE survey. They are indicated in red in the figure and increase the SDSS imaging footprint by roughly 2000 deg^2 over DR6. Note that many of these cross the Galactic plane (indicated by the sinuous line crossing the figure). Unlike DR6, the union of the Legacy and SEGUE data are now available in a single database in CAS in DR7.

[104] http://cas.sdss.org/astro
[105] http://das.sdss.org

Table 1
Coverage and Contents of DR7

Imaging	
Imaging area in CAS	11,663 deg^2
Imaging catalog in CAS	357 million unique objects
Legacy footprint area	8423 deg^2
	(7646 deg^2 in North Galactic Cap)
Legacy imaging catalog	230 million unique objects
	585 million entries (including duplicates)
SEGUE footprint area, available in DAS[a]	3500 deg^2 (more than double DR6)
SEGUE footprint area, available in CAS	3240 deg^2
SEGUE imaging catalog	127 million unique objects
M31, Perseus, Sagittarius scan area	∼46 deg^2
Southern Equatorial Stripe with >70 repeat scans	∼250 deg^2
Commissioning ("Orion") data	832 deg^2
Spectroscopy	
Spectroscopic footprint area	9380 deg^2
Legacy	8032 deg^2
SEGUE	1348 deg^2
Total number of plate observations (640 fibers each)	2564
Legacy Survey plates	1802
SEGUE and special plates	676
Repeat observations of plates	86
Total number of spectra[b]	1,630,960
Galaxies	929,555
Quasars	121,363
Stars	464,261
Sky	97,398
Unclassifiable	28,383
Spectra after removing skies and duplicates	1,440,961

Notes.
[a] Includes regions of high stellar density, where the photometry is likely to be poor. See the text for details. This area also includes some regions of overlap.
[b] Spectral classifications from the `spectro1d` code; numbers include duplicates.

These data have been recalibrated using ubercalibration (Padmanabhan et al. 2008) using the overlap between adjacent scans; the resulting photometry is now the default photometry found in the CAS. We also make available the original photometry calibrated by the auxiliary Photometric Telescope (Tucker et al. 2006). The ubercalibration solution was regenerated using all the imaging data, but the changes are tiny from the ubercalibration results published in DR6: 0.001 mag rms in $griz$ and 0.003 mag in u. The ubercalibrated photometry zero points are defined to be the same as that measured from the Photometric Telescope.

The green and blue patches indicate supplementary imaging stripes, which contain scans over M31 or in its halo, through the center of the Perseus cluster of galaxies, over the low-latitude globular cluster M71, near the South Galactic Pole, along the orbit of the Sagittarius Tidal Stream, and through the star-forming regions of Orion (Finkbeiner et al. 2004). In addition, there are a number of scans at angles perpendicular, or at an oblique angle, to the regular Legacy or SEGUE imaging stripes. These scans are used in the ubercalibration procedure to tie the zero points of the stripes together and to determine the flat fields.

The lower panel in Figure 1 shows the coverage of spectroscopy in DR7; the light gray area shows the increment in the Legacy Survey over DR6. Most importantly, the gap cutting the North Galactic Cap in two pieces in previous data releases has been closed; we now have complete spectroscopy of our principal galaxy and quasar targets over a contiguous area of roughly 7500 deg^2. An additional dozen plates were observed to fill holes in the nominally contiguous regions in DR6. Adding in the three stripes in the Southern Galactic Cap, the Legacy spectroscopy footprint is 8032 deg^2, a 26% increment over DR6.

In addition, spectroscopy was carried out using a series of target selection algorithms designed to find stars of a wide variety of types as part of the SEGUE project (DR6 paper; Yanny et al. 2009). These targets were drawn from both the SEGUE and Legacy imaging, and are shown in red in the lower panel of Figure 1. As some of these are lost in the density of Legacy spectra, we show the distribution of SEGUE and other non-Legacy spectra in Galactic coordinates in Figure 2.

Finally, as described in Yanny et al. (2009), we carried out spectroscopy of stars in 12 open and globular clusters to calibrate the measurements of stellar parameters in SEGUE (Lee et al. 2008a, 2008b). Many of these clusters are sufficiently close that the giant branches are brighter than the photometric saturation limit of SDSS, so the targets for these plates were selected from the literature. Indeed, the spectrographs would saturate as well with our standard 15 minute exposures, so these observations had individual exposure times as short as 1 or 2 minutes. Without proper flux calibrators or exposure of bright sky lines to set the zero point of the wavelength scale, the spectrophotometry and wavelength calibration of the spectra on these plates are often quite inferior to that of the main survey, and these plates are available only in the DAS, not the CAS.

As described in more detail below, the 2.°5 stripe centered on the celestial equator was imaged multiple times through-

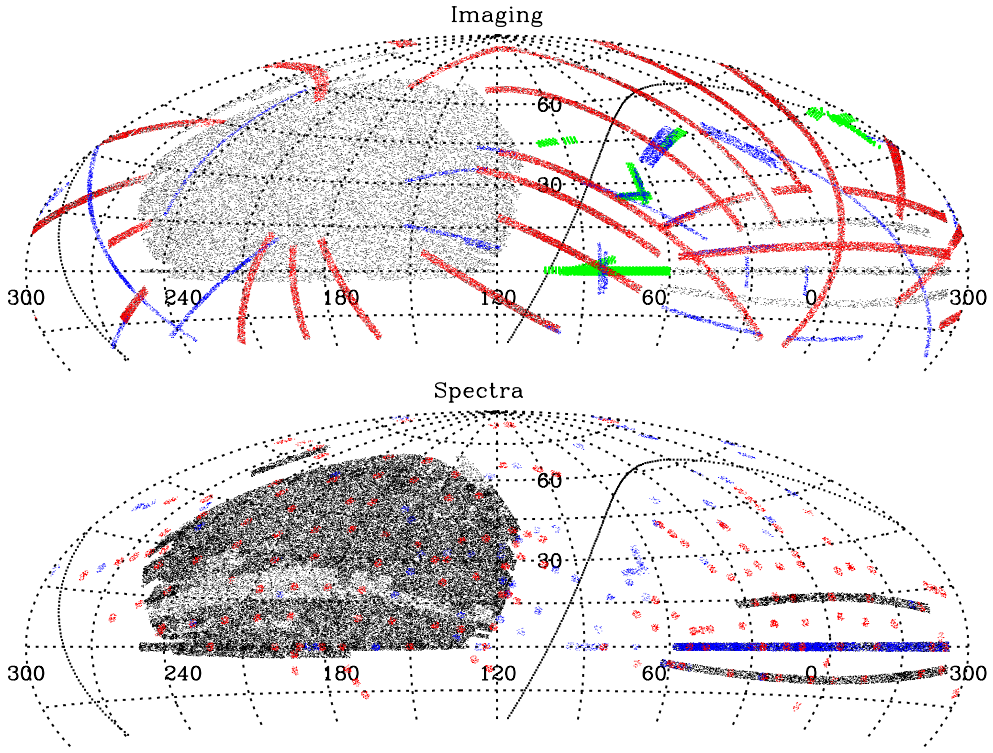

Figure 1. Distribution on the sky of the data included in DR7 (upper panel: imaging; lower panel: spectra), shown in an Aitoff equal-area projection in J2000 Equatorial Coordinates. The Galactic plane is the sinuous line that goes through each panel. The center of each panel is at $\alpha = 120^\circ \equiv 8^{\rm h}$, and the plots cut off at $\delta = -25^\circ$, below which the SDSS did not extend. The Legacy imaging survey covers the contiguous area of the Northern Galactic Cap (centered roughly at $\alpha = 200^\circ$, $\delta = 30^\circ$), as well as three stripes (each of width 2.°5) in the Southern Galactic Cap. In addition, several stripes (indicated in blue in the imaging data) are auxiliary imaging data, while the SEGUE imaging scans are indicated in red. The green scans are additional runs as described in Finkbeiner et al. (2004). In the spectroscopy panel, the lighter regions indicate that area in the Northern Galactic Cap which is new to DR7; note that the Northern Galactic Cap is now contiguous. Red points indicate SEGUE plates and blue points indicate other non-Legacy plates (mostly as described in the DR4 paper).

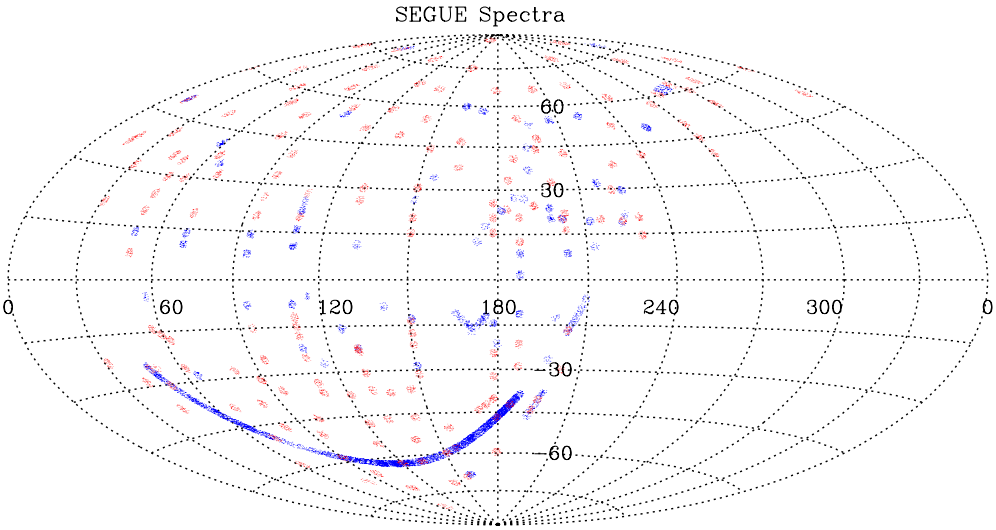

Figure 2. Distribution on the sky of SEGUE (red) and other non-Legacy (blue) spectroscopic observations, here plotted in Galactic coordinates. The contiguous blue stripe across the bottom is Stripe 82, along the celestial equator. As described in the DR4 paper, Stripe 82 includes extensive spectroscopy of a number of different types of targets outside the Legacy Survey.

out SDSS and SDSS-II. Each 2.°5 wide stripe is observed by a pair of offset *strips* to cover the full width (York et al. 2000); the coverage of the two strips of Stripe 82 is shown in Figure 3. The data are shown both for the subset of data included in a deep co-addition (the lower set of curves; Section 3.3) and all scans, including those taken under nonideal conditions for the supernova survey (Section 3.2; Frieman et al. 2008).

3. ADDITIONAL IMAGING PRODUCTS AND DATABASES

3.1. The Runs Database

The SDSS imaging survey was primarily designed to give a single pass across the sky, thus in the CAS, each photometric measurement is flagged either `Primary` or `Secondary`. `Primary` objects designate a unique set of detections (i.e.,

without duplicates) using the geometric boundaries of survey stripes.[106] The set of `Secondary` objects includes repeat observations of the same object in overlapping strips and stripes. `Primary` objects are associated with a run and field which is the primary source of imaging data at that position. In DR7, the union of the Legacy and SEGUE footprints serves as the `Primary` footprint; a quantity `inLegacy` in the `fieldQA` table in CAS indicates those objects which lie within the original Legacy Northern Galactic Cap Survey ellipse, as defined in York et al. (2000). Legacy imaging can also be distinguished by the stripe number for each run; Stripes 9–44, 76, 82, and 86 are in the Legacy Survey, all others are SEGUE stripes or other miscellaneous pieces of sky (Figure 1).

While resolving the sky into a seamless `Primary` region of unique detections of objects is ideal for many science queries, it is sometimes convenient to query data by run without regard to the way the survey resolves overlaps and imposes the boundaries of the edge of the survey. These boundaries are restricted to matched pairs of North and South strips in the main DR7 CAS. Therefore in many runs, several fields at the beginning or end which do not have a match in the corresponding other strip are not included in the main CAS. Thus, we have now made available a separate `runs` database within the CAS, which includes *all* fields in all runs, and which allows one to query objects by which run they are imaged in.

The `runs` database contains 530 complete runs from SDSS-I and SDSS-II, where `Primary` is set strictly based on geometric limits within each scan, regardless of overlapping runs or stripes. The `runs` database also contains several scans outside the regular DR7 Legacy or SEGUE footprints. For example, Stripe 205 is covered by runs 4334, 4516, 6751, and 6794, and follows the Sagittarius Stream, which is in three pieces, the first running from $(\alpha, \delta) = (240^\circ, -15^\circ)$ to $(200^\circ, +10^\circ)$, the second centered at $(135^\circ, +35^\circ)$, and the last (overlapping several other runs) which ends at $(45^\circ, +10^\circ)$.

3.2. The Stripe 82 Database

The SDSS stripe along the celestial equator in the Southern Galactic Cap ("Stripe 82") was imaged multiple times in the Fall months. This was first carried out to allow the data to be stacked to reach fainter magnitudes, and through Fall 2004, these data were taken only under optimal seeing, sky brightness, and photometric conditions (i.e., the conditions required for imaging in the Legacy Survey; York et al. 2000). There were 84 such runs made public in previous data releases. In Fall 2005, 2006, and 2007, 219 additional imaging runs were taken on Stripe 82 as part of the SDSS supernova survey (Frieman et al. 2008), often under less optimal conditions: poor seeing, bright moonlight, and/or nonphotometric conditions. These data have been photometrically calibrated following the prescription of Bramich et al. (2008), whereby the photometry of bright stars is tied to that of photometric data on a field-by-field basis (see Ivezić et al. 2007 for a similar approach). Bramich et al. solved for photometric offsets both parallel and perpendicular to the scan direction in data from a given CCD; we found that the term perpendicular to the scan direction added little, and we did not include it here. As Bramich et al. (2008) show, the resulting photometric calibration is good to 0.02 mag at the bright end in up to 1 mag of atmospheric extinction. Of course, under nonoptimal conditions, these data will not necessarily reach as deep as normal survey images.

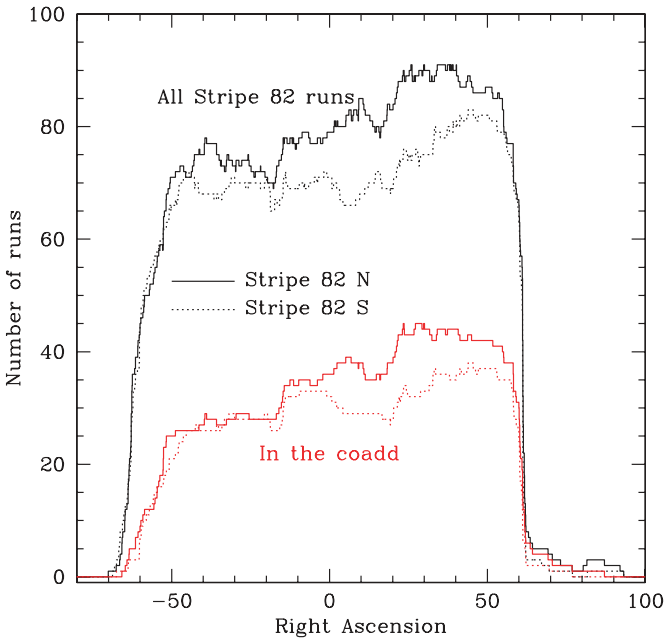

Figure 3. Stripe 82, the equatorial stripe in the South Galactic Cap, has been imaged multiple times. The lower pair of curves shows the number of scans covering a given right ascension in the North and South strips that are included in the co-addition (mostly data taken through 2005). In addition, Stripe 82 has been covered many more times as part of a comprehensive survey for $0.05 < z < 0.35$ supernovae, although often in conditions of poor seeing, bright moon, and/or clouds; the total numbers of scans at each right ascension in the North and South strips are indicated in the upper pair of curves. All these data have been flux calibrated, as discussed in the text, and are available (together with the co-add itself) in the `stripe82` database.

(A color version of this figure is available in the online journal.)

SDSS judges photometricity of a given night by monitoring fluctuations in the night sky measured by a wide area infrared camera (the "cloud camera") sensitive at 10 μm, where clouds are emissive (Hogg et al. 2001). If the sky fluctuations are small and constant, then the night is photometric. Clouds cause the fluctuations to increase. Plots of cloud cover and seeing for most nights on which Stripe 82 was observed are available as part of the DR7 web documentation listing all Stripe 82 scans. In addition, for those runs which the cloud camera indicated as nonphotometric, we examined the fluctuations in the zero point for each CCD in the camera as a function of time using the photometric calibration procedure of Bramich et al. (2008). These zero-point values are available in the CAS; rms variations of more than 0.1 mag are an indication of considerable variable cloud cover, and a value of more than 1 mag suggests that the approximate calibration procedure of Bramich et al. (2008) breaks down, and the resulting photometry should be regarded with caution. All 303 runs covering Stripe 82 are made available as part of the `Stripe82` database, which is structured like the `runs` database.

3.3. Going Deep on Stripe 82

We have carried out a co-addition of the repeat imaging scans on Stripe 82 taken through Fall 2005 under the best conditions (see below). The co-addition includes a total of 122 runs, covering any given piece of the >250 deg^2 area between 20 and 40 times (Figure 3), and the results are made available in the `Stripe82` database as well. The co-addition runs are designated

[106] See http://www.sdss.org/dr7/algorithms/resolve.html for a detailed explanation.

100006 (South strip) and 200006 (North strip), respectively in the DAS, and 106 and 206 in the CAS.

The co-addition is described in detail in J. Annis et al. (2009, in preparation); see also Jiang et al. (2008). From the list of runs on Stripe 82 taken through the Fall 2005 season, all fields with seeing in the r band worse than $2''$ FWHM, r-band sky brightness brighter than 19.5 mag in 1 square arcsecond, or whose photometric correction à la Bramich et al. (2008; see above) was greater than 0.2 mag were excised; this rejected 32% of the available data. The individual runs were remapped onto a uniform astrometric coordinate system. Interpolated pixels in each individual run (e.g., for bad columns, bleed trails, and cosmic rays) were masked in the co-addition process. The sky was subtracted from each frame, and the images co-added with weights for each frame proportional to the transparency and inversely proportional to the square of sky noise and seeing on each frame. Strongly discrepant pixels were clipped in the co-addition. The effective seeing FWHM is $\sim 1\farcs2$ (for the southern strip of the stripe) and $\sim 1\farcs3$ (for the northern strip).

The resulting co-added images were run through the SDSS photometric pipeline, yielding the catalog made available in the `Stripe82` database. Rather than deriving the point-spread function (PSF) from scratch, we synthesized the PSF at each point in the sky by taking the suitably weighted sum of the PSFs output by the SDSS photometric pipeline from each of the individual runs.

Color–color diagrams of stars and counts of stars and galaxies as a function of magnitude demonstrate that the photometry reaches roughly 2 mag fainter than single SDSS scans, similar to what is expected given the number of runs in the co-add. We have found that star–galaxy separation is improved over that in the single scans, in that the cut can be made closer to the stellar locus. In the main survey, objects with $m_{\rm PSF} - m_{\rm model} > 0.145$ are flagged as galaxies in a given band. However, the stellar peak in the PSF − model magnitude difference distribution in the co-add is much narrower, allowing objects with $m_{\rm PSF} - m_{\rm model} > 0.03$ in r to be flagged as galaxies.

The co-addition does not properly propagate information on saturated pixels in individual runs, and therefore the photometry of objects brighter than roughly $r = 15.5$ is suspect. Unfortunately, there is no processing flag that one can use to identify such data; we recommend a simple magnitude cut.

The SDSS photometry is quoted in terms of asinh magnitudes, as described by Lupton et al. (1999), whereby the logarithmic magnitude scale transitions to a linear scale in flux density f at low S/N:

$$m = -\frac{2.5}{\ln 10}\left[\operatorname{asinh}\left(\frac{f/f_0}{2b}\right) + \ln(b)\right]. \quad (1)$$

The magnitude at which this transition occurs is set by the quantity b, which is roughly the fractional noise in the sky in a PSF aperture in $1''$ seeing (EDR paper). Here $f_0 = 3631$ Jy, the zero point of the AB flux scale. The quantity b for the co-addition is given in Table 2, along with the asinh magnitude associated with a zero-flux object. Compare with the equivalent numbers for the main survey, given in Table 21 of the EDR paper. Table 2 also lists the flux corresponding to $10 f_0 b$, above which the asinh magnitude and the traditional logarithmic magnitude differ by less than 1% in flux.

As with the main survey, it is important to use the various processing flags output by the photometric pipeline (e.g., as recommended by Richards et al. 2002) to reject spurious objects, and to select objects with reliable photometry.

Table 2
Asinh Magnitude Softening Parameters for the Co-Addition

Band	b	Zero-Flux Magnitude ($m(f/f_0 = 0)$)	m ($f/f_0 = 10b$)
u	1.0×10^{-11}	27.50	24.99
g	0.43×10^{-11}	28.42	25.91
r	0.81×10^{-11}	27.72	25.22
i	1.4×10^{-11}	27.13	24.62
z	3.7×10^{-11}	26.08	23.57

4. IMPROVEMENTS IN PROCESSING OF IMAGING DATA

4.1. New Reductions of SEGUE Imaging Data and Crowded Fields

As was noted in the DR6 paper, the SDSS imaging pipeline (`photo`) was designed to analyze data at high Galactic latitudes, and is not optimized to handle very crowded fields. The Legacy Survey is restricted to high latitudes, and `photo` performs adequately throughout the Legacy footprint. However, at lower latitudes, when the density of stars brighter than $r = 21$ grows above 5000 deg^{-2}, the pipeline is known to fail, as it is unable to find sufficiently isolated stars to measure an accurate PSF, and the deblender does poorly with overly crowded images. Many of the SEGUE scans probe these low latitudes (Figure 1), and we therefore adapted an alternative stellar photometry code called `PSPhot` developed by the Pan-STARRS team (Kaiser et al. 2002; Magnier 2006) to be used for these runs. In brief, we first run this code, and then run `photo` using the list of objects detected by `PSPhot` as input to help `photo`'s object finder in crowded regions. This approach thus provides two sets of photometry at low latitudes.

Like, e.g., DAOPHOT (Stetson 1987), `PSPhot` begins with the assumption that every object is unresolved, and therefore does a better job than `photo` in crowded stellar regions. It uses an analytical model based on Gaussians to describe the basic PSF shape, with parameters which may vary across the field of the image to follow the PSF variations. It also uses a pixel-based representation of the residuals between the PSF objects and the analytical model, which is also allowed to vary across each field. Candidate PSF stars are selected from the collection of bright objects in the frame by searching for a tight clump in the distribution of second moments. After rejecting outliers, the PSF fit parameters are used to constrain the spatial variations in the PSF model.

Unlike `photo`, `PSPhot` processes each frame separately (without any requirement of continuity of PSF estimation across frame boundaries), and each filter separately (without any requirement that the lists of objects between the separate filters agree). The pipeline outputs positions and PSF magnitudes (and errors) for each detected object; the results are found in the `PsObjAll` table in the CAS. The resulting photometry is then matched between filters using a $1''$ matching radius. While the estimated PSF errors output by `photo` include a term from the uncertainty in the PSF fitting, this component is not included in the errors reported by `PSPhot`.

We then run `photo`, asking it to carry out photometry at the position of each object detected by `PSPhot`, in addition to the positions of objects `photo` itself detects. This allows `photo` to do a much better job of distinguishing individual objects in crowded regions. In addition, the pipeline is fine tuned to less aggressively look for overlap between adjacent objects, and not to give up as soon as it does at high latitude when faced with

deblending large numbers of objects. We describe below how the photometry directly out of PSPhot, and that from photo, compare.

The SDSS PSF photometry had an offset applied to it to make it agree with aperture photometry of bright stars within a radius of 7.″43; this large-aperture photometry was in fact what was used by ubercalibration to tie all the photometry together (Padmanabhan et al. 2008). In crowded regions, finding sufficiently isolated stars to measure aperture photometry becomes difficult. PSPhot photometry was forced to agree with these large-aperture magnitudes for bright stars; this was done in practice by determining, for each CCD in the imaging camera for each run, the average aperture correction needed to put the two on the same system, using stars at Galactic latitude $|b| > 15°$, where crowding effects are less severe.

If any part of a SEGUE imaging run extended to $|b| < 25°$, the entire run was processed through the photo and PSPhot code. This sample includes most (but not all) of the SEGUE imaging runs. These PSPhot+photo processed runs, designated with rerun = 648 in the DR7 CAS and DAS, are declared the Best reduction of these SEGUE runs. There is also an inferior Target version of these SEGUE runs which was used to design SEGUE spectroscopic plates; it is based on photo alone, as the PSPhot code was unavailable at the time the plates were designed. The Target reductions have rerun = 40 and are segregated to the SEGUETARGDR7 database.

This processing revealed a problem with photo. In crowded regions, one cannot find sufficiently isolated stars to measure counts through such a large aperture, and in practice, the code corrected PSF magnitudes to an aperture photometry radius of 3.″00 instead, whenever any part of a given run dipped below $|b| = 8°$. Thus, the aperture correction was underestimated, typically by 0.03–0.06 mag, depending on the seeing. This was not a problem for any of the Legacy imaging scans, but is very much an issue for the SEGUE runs. Fortunately, there is a strong correlation, in a given detector, between the aperture correction from a 3.″00 aperture to a 7.″43 aperture (as measured on high-latitude fields), and the seeing. We therefore applied this correction after the fact to the photo PSF, de Vaucouleurs, exponential, and model magnitudes for all SEGUE runs affected by this problem. This was carried out before ubercalibration, so these runs are photometrically calibrated in a consistent way.

4.2. *Comparison of* photo *and* PSPhot *Photometry*

The quality of the photometry produced by PSPhot and by photo with the PSPhot-detected objects as input, was evaluated by comparing the magnitudes computed by the two methods. Within each field, we calculated the median of the difference of PSF magnitudes for stars with $14 < u, g, r, i, z < 20$. This median difference had an rms of 0.014 mag. Fields with a difference greater than 0.08 mag are suspect, and further investigation is needed to determine which of the two pipelines might be at fault. We followed McGehee et al. (2005) to measure reddening-free colors of the same stars that track the stellar locus:

$$Q_{gri} = (g - r) - E_{gri}\,(r - i), \quad (2)$$

$$Q_{riz} = (r - i) - E_{riz}\,(i - z), \quad (3)$$

where $E_{gri} = 1.582$ and $E_{riz} = 0.987$. These are normalized to equal zero at high Galactic latitude (note that these colors do *not* include the *u* band).

Median Q_{gri} and Q_{riz} colors in each field were computed for objects identified as stars in each field, and satisfying

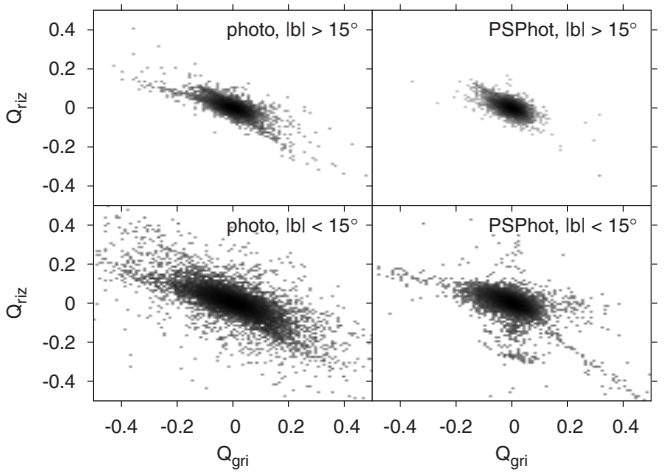

Figure 4. Distribution of median Q_{gri}- and Q_{riz}-parameters measuring the position of the stellar locus within each field for the photo (left) and PSPhot (right) photometric pipelines; zero values are indicative of uniform photometry. Within the Galactic plane (lower panels), the PSPhot values are more concentrated, but contain a higher number of systematic departures from the main locus. The PSPhot code in fact gives a tighter locus at high latitudes as well (upper panels). Histogram equalization of the grayscale was used to emphasize low-density regions.

magnitude and color cuts as follows: $14 < (u, g, r, i, z) < 20$, $0.5 < (u-g) < 1.9, 0.0 < (g-r) < 1.2, -0.2 < (r-i) < 0.8$, and $-0.2 < (i - z) < 0.6$. The *Q*-parameters were found to be lower by up to 0.1 mag at low Galactic latitudes; to remove this effect, we fitted a model of a constant plus Lorentzian to the median *Q* values as a function of Galactic latitude, and subtracted it. The distributions of the Q_{gri} and Q_{riz} values for both photo and PSPhot are compared as density plots in Figure 4. From Equation 2, photometric errors in a single filter manifest themselves differently: δg as a shift in Q_{gri}, δr as a line with slope $dQ_{riz}/dQ_{gri} = -1/(1 + E_{gri}) = -0.35$, δi as a line with slope $dQ_{riz}/dQ_{gri} = -(1 + E_{riz})/E_{gri} = -1.07$, and δz as a shift in Q_{riz}.

The photo data in a given field were flagged as bad when either $|Q_{gri}|$ or $|Q_{riz}| > 0.12$ mag ($\gtrsim 5\sigma$) as measured from photo magnitudes, and similarly for the PSPhot outputs. Of course, a field could be flagged as bad in both sets of outputs. By this criterion, about 2% of the fields processed with PSPhot were flagged bad based on the photo outputs, and 3.6% were bad based on PSPhot photometry. The vast majority of the flagged fields are within 15° of the Galactic plane, and essentially all the fields in which the median difference between photo and PSPhot photometry was greater than 0.08 mag in a given band were flagged as bad by the *Q* criteria. This flag and the Q_{gri} and Q_{riz} quantities themselves can be found in the fieldQA table in the CAS.

Although more fields are flagged based on the PSPhot outputs, the PSPhot scatter in Figure 4 is tighter at both high and low Galactic latitudes than for photo. The PSPhot stellar photometry is therefore preferred for studies of the stellar locus (we have not fully assessed its robustness to outliers), but comes with the caveat that fields flagged bad should be identified in the fieldQA table and be culled.

An alternative check of SDSS photometry in dense stellar fields was carried out by An et al. (2008), who reduced the SDSS imaging data for crowded open and globular cluster fields using the DAOPHOT/ALLFRAME suite of programs (Stetson 1987, 1994). At a stellar density of ~400 stars deg^{-2} with $r < 20$,

they found ∼2% rms variations in the difference between photo and DAOPHOT magnitudes in the scanning direction in all five bandpasses (see their Figure 3). The systematic structures are likely due to imperfect modeling of the PSFs in photo, given that DAOPHOT magnitudes exhibit no such large variations with respect to aperture photometry. In other words, the PSF variations were too rapid for the photo pipeline to follow over a timescale covered approximately by one field (≈ 10′ or ≈54 s in time).

An et al. (2008) further examined the accuracy of photo magnitudes in semicrowded fields using three open clusters in their sample. Stellar densities in these fields were as much as ∼10 times higher than those in the high Galactic latitude fields, but photo recovered ∼80%–90% of stellar objects in the DAOPHOT/ALLFRAME catalog. An et al. (2008) found that these fields have only marginally stronger spatial variations in photo magnitudes than those at lower stellar densities.

4.3. Further Assessments of Imaging Quality

Section 4.6 of the EDR paper describes a series of flags available in the database to assess the quality of each field in the imaging data; this includes information on whether the data in a given field meet survey requirements on seeing and sky brightness. We have added additional criteria to assess the quality of each field. The CAS table called fieldQA includes a flag called ProblemChar associated with each field, which is set when:

1. The median of the telescope focus over three frames moved more than 60 μm, indicating a problem with the automated focus of the telescope (Gunn et al. 2006; Problemchar = "f").
2. The rotator angle moved more than 25″ between adjacent fields (corresponding to a 0.″55 image shift at the edge of the camera) (ProblemChar="r").
3. The astrometric solution shifted by more than 4 pixels (1.″6) from a smooth interpolation between adjacent fields (ProblemChar="a").
4. Miscellaneous other problems, including voltage problems in the camera, and lights left on in the telescope; this was triggered in only two imaging runs (ProblemChar="s").

We flag all fields with these problems in any of the five bandpasses. Because the imaging observations are done in drift-scan mode (Gunn et al. 1998), different areas of the sky are observed simultaneously in each bandpass and referenced to the field number of the *r*-band observation. Thus in the case of focus problems, we mark the 11 fields preceding, and the three fields following the field in question in all camera columns in the run as bad. For the rotator and astrometric shift problems, we similarly mark the nine preceding and the one following field as bad. Only 0.3% of all fields in the CAS are marked with one of these problems (the majority of which are due to focus problems); these flags should be consulted when examining the reliability of the photometry in a given area of sky.

4.4. Astrometric Recalibration

Early SDSS imaging runs were astrometrically calibrated against Tycho-2 (Høg et al. 2000), which yielded statistical errors per coordinate for bright stars ($r < 20$) of approximately 75 mas and systematic errors of 20–30 mas. Later runs were calibrated against preliminary versions of the USNO CCD Astrograph Catalog (UCAC; Zacharias et al. 2000), which yielded improved statistical errors per coordinate of approximately 45 mas, with systematic errors of 20–30 mas (Pier et al. 2003). Proper motions were not available for the preliminary versions of UCAC. Since the typical epoch difference between the SDSS and UCAC observations is a few years and the typical proper motion of UCAC stars is a few mas year^{-1}, this introduces an additional roughly 10 mas of systematic error in the positions due to the uncorrected proper motions of the calibrating stars.

All of DR7 has been recalibrated astrometrically against the Second Data Release of UCAC (UCAC2; Zacharias et al. 2004). While the systematic errors for UCAC2 are not yet well characterized, they are thought to be less than 20 mas (N. Zacharias 2008, private communication). UCAC2 also includes proper motions for stars with $\delta < +41^\circ$. For stars at higher declination, proper motions from the SDSS+USNO-B catalog (Munn et al. 2004) have been merged with the UCAC2 positions. With these improvements, all DR7 astrometry has statistical errors per coordinate for bright stars of approximately 45 mas, with systematic errors of less than 20 mas. The mean differences per run between the old and new calibrations is a function of position on the sky, with typical absolute mean differences of 0–40 mas. The rms differences are of order 10–40 mas for runs previously reduced against UCAC, and 40–80 mas for runs previously reduced against Tycho-2, consistent with what we would expect given the errors in the reductions.

Note that the formal SDSS names of objects in the CAS are of the form SDSS Jhhmmss.ss±ddmmss.s. Because of the subtle changes in the astrometry, these names will be slightly different for many objects between DR6 and DR7. The user should be aware of this in comparing objects between DR6 and DR7.

The CAS includes proper motions for objects derived by combining SDSS astrometry with USNO-B positions, recalibrated against SDSS (Munn et al. 2004). These are given in the ProperMotions table in the CAS.[107] An error was discovered in the proper motion code in Data Releases 3 through 6, which causes smoothly varying systematic errors, in the proper motion in right ascension only, of typically 1–2 mas year^{-1} (see Munn et al. 2007 for a full description of the problem and its effects). This error has been corrected in DR7, thus any use of proper motions should use the DR7 CAS.

4.5. SEGUE Target Selection

Several of the SEGUE target selection algorithms evolved during the course of SDSS-II. The most significant changes occurred to the K-giant algorithm, as it was realized that good color-based luminosity separation could be done only for the very reddest ($g - r > 1.1$) giant candidates by their deviation from the main-sequence locus in the *ugr* color diagram; this of course requires accurate *u*-band photometry. Early K-giant target selection included stars with $(g-r)_0$ (where the subscript refers to values after correcting for SFD Galactic extinction) as blue as 0.35. The final selection chose stars with $(g - r)_0$ between 0.5 and 1.2 and was restricted to $g_0 < 18.5$; this gives much cleaner samples of K giants (Yanny et al. 2009).

In order to allow users to analyze completeness and efficiency of SEGUE stellar target selection samples, the latest (v4.6) version of the algorithms (Yanny et al. 2009) was applied to *all* stellar objects in the imaging catalog which had $g < 21$ or $z < 21$, over the entire sky. The appropriate bits were propagated into the SEGUEPrimTarget and SEGUESecTarget fields of the

[107] This table was called USNOB in the DR3 and DR4 versions of the CAS.

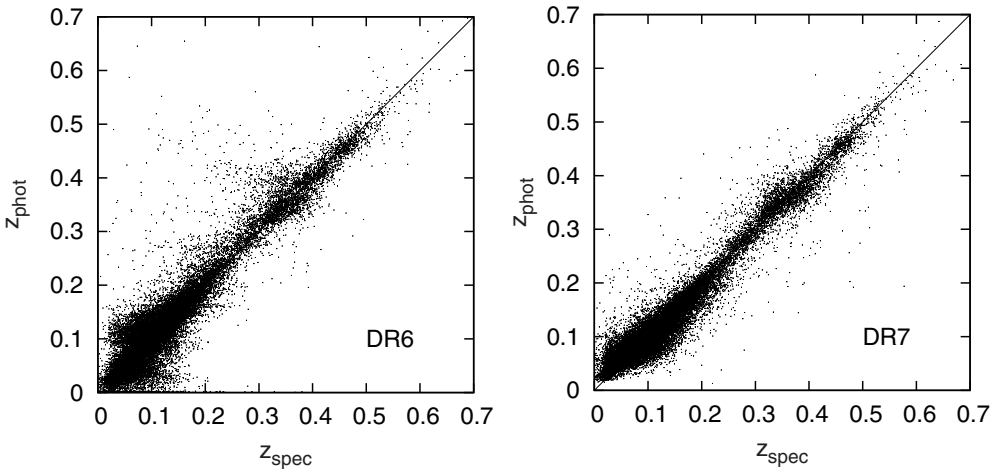

Figure 5. Template-based estimated redshifts vs. the true spectroscopic redshifts for a random sample of 30,000 galaxies with redshifts from SDSS. The estimated values calculated with the old (DR6) method have significantly larger scatter and more outliers than the ones with the new hybrid (DR7) technique. Note that the sample is dominated by red galaxies (whose photometric redshifts are intrinsically easier to measure) at $z > 0.2$.

`photoObjAll` table of the DR7 CAS. A description of the bits and the target selection algorithms is available in Yanny et al. (2009).

4.6. Photometric Redshifts

As described in the DR5 paper, the SDSS makes available the results of two different photometric redshift determinations for galaxies, one based on neural nets and the other based on a template-fitting approach. With DR7, we include improvements to both, as we now describe.

4.6.1. Photometric Redshifts with Neural Nets

The neural net solutions for photometric redshifts and their errors (listed as `Photoz2` in the CAS, and described in Oyaizu et al. 2008) have not changed since DR6, and do not use the ubercalibrated magnitudes. However, we now provide a value-added catalog containing the redshift probability distribution for each galaxy, $p(z)$, calculated using the weights method presented in Cunha et al. (2008). The $p(z)$ for each galaxy in the catalog is the weighted distribution of the spectroscopic redshifts of the 100 nearest training-set galaxies in the space of dereddened model colors and r magnitude. For the $p(z)$ calculation, we also added the zCOSMOS (Lilly et al. 2007) and DEEP2-EGS (Davis et al. 2007) galaxies to the spectroscopic training set used for the `Photoz2` solution.

Cunha et al. (2008) showed that summing the $p(z)$ for a sample of galaxies yields a better estimation of their true redshift distribution than that of the individual photometric redshifts. Mandelbaum et al. (2008) found that this gives significantly smaller photometric lensing calibration bias than the use of a single photometric redshift estimate for each galaxy.

4.6.2. Photometric Redshifts: A New Hybrid Technique

With DR7, we have made substantial improvements in the other photometric redshift code (`Photoz`), using a hybrid method that combines the template-fitting approach of Csabai et al. (2003; i.e., the approach used in DR5 and DR6) and an empirical calibration using objects with both observed colors and spectroscopic redshifts. We summarize the method briefly here, with details to follow in a paper in preparation.

The spectroscopic sample of SDSS contains over 900,000 spectroscopically confirmed galaxies, and the combination of the main sample (Strauss et al. 2002), the LRG sample (Eisenstein et al. 2001) and special plates targeted at fainter blue galaxies (DR4 paper) more or less cover the whole color region in which galaxies lie to the depths of SDSS. Thus, we use the DR7 spectroscopic set as a reference set for redshift estimation without any additional data from synthetic spectra.

The estimation method uses a k-d tree (following Csabai et al. 2007) to search in the ubercalibrated $u-g$, $g-r$, $r-i$, and $i-z$ color space for the 100 nearest neighbors of every object in the estimation set (i.e., the galaxies for which we want to estimate redshift) and then estimates redshift by fitting a local hyperplane to these points, after rejecting outliers. If an object lies outside the bounding box of the 100 nearest neighbors in color space, the photometric redshift is less reliable, and the object is flagged accordingly. We use template fitting to estimate the K-correction, distance modulus, absolute magnitudes, rest frame colors, and spectral type. We search for the best match of the measured colors and the synthetic colors calculated from repaired (Budavári et al. 2000) empirical template spectra at the redshift given by the local nearest neighbor fit.

We have found that the mean deviations of the redshifts from the best-fit hyperplane is a good estimate of the error. That, together with the flag indicating whether an object lies outside the bounding box of its neighbors, and the difference between the estimated photometric redshift and the average redshift of its neighbors, can be used to select objects with reliable photometric redshift values.

The rms error of the redshift estimation for the reference set decreases from 0.044 in DR6 to 0.025 in DR7 with this improved algorithm (Figure 5). Iteratively removing the outliers beyond 3σ gives rms errors of 0.028 and 0.020 for the old and new methods, respectively. In addition, the reliability of the quoted errors is much higher.

4.7. SDSS Filter Response Functions

The response functions of the SDSS imager as a function of wavelength have been monitored throughout the survey. The $griz$ responses were stable over time, although very small seasonal (i.e., temperature) variations were observed, at a level well below our typical photometric errors. However, we have found a relatively large change in both the amplitude and shape of the u-band response, which is likely due to

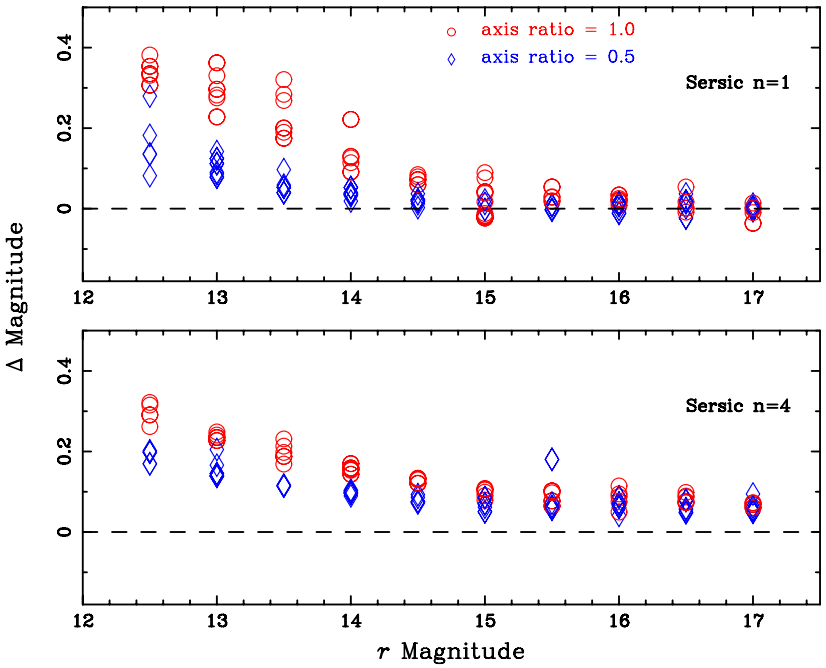

Figure 6. Difference between measured model and true *r*-band magnitudes of a series of simulated galaxies with Sérsic indices of 1 (disk galaxies; upper panel) and 4 (elliptical galaxies; lower panel). These galaxies followed the magnitude–effective radius relation observed in the SDSS Value-Added Galaxy Catalog (Blanton et al. 2005), and were either circularly symmetric (circles) or had an axis ratio of 0.5 (diamonds). They were added to random areas of real high-latitude fields and run through `photo`. The simulated elliptical galaxies show a systematic offset even at the faint end; this is due to the fact that the `photo` model magnitude code assumes a truncation beyond seven scale lengths, while the "true" magnitude has no such truncation. This is a 0.05 mag effect.

(A color version of this figure is available in the online journal.)

a degradation of the UV-enhanced coating of the *u*-band CCD. This change in instrumental zero point is effectively corrected by the photometric calibration for objects near the mean color of the standard stars, and, in fact, the repeat photometry of stars in Stripe 82 is stable with time for stars with $-0.5 < g - r < 1.5$ (Bramich et al. 2008; Ivezić et al. 2007). However, the observed response changes involve a roughly 30 Å redward shift in the effective wavelengths of the *u* filters over the lifetime of the survey, so one would expect significant changes in the measured colors of objects of extreme color over the period, and this is being investigated. M. Doi et al. (2009, in preparation) will summarize the filter characteristics in full, including column-to-column variations within the camera and the changes with time.

5. PHOTOMETRY OF BRIGHT GALAXIES

As described in the DR6 paper and Mandelbaum et al. (2005), systematic errors in the estimation of the sky near bright ($r < 16$) galaxies cause their fluxes and scale sizes to be underestimated and the number of neighboring objects to be suppressed. Indeed, a number of authors (Lauer et al. 2007, Bernardi et al. 2007, Lisker et al. 2007) have pointed out systematic errors in SDSS galaxy photometry at the bright end. In the DR6 paper, this effect was quantified by adding simulated galaxies to the SDSS images using a code described in Masjedi et al. (2006). These simulations found that the *r*-band brightness of galaxies was underestimated by as much as 0.8 mag for a 12th magnitude galaxy with Sérsic index, $n = 1$ (an "exponential," or disk galaxy). For $n = 4$ galaxies ("de Vaucouleurs," or elliptical galaxies), the effect was less pronounced, with a brightness underestimate of less than 0.6 mag.

However, the simulations shown in the DR6 paper used an incorrect relation between galaxy size and magnitude, in the sense that they overestimated the extent of the problem for the typical galaxy. Using instead the relationships between apparent magnitude and half-light radius measured for SDSS bulge and disk galaxies (Blanton et al. 2003), we repeated the exercise: we simulated pure $n = 1$ and $n = 4$ galaxies with axis ratios b/a of 0.5 and 1, and added them to real *r*-band SDSS images. We ran the results through `photo` and compared their measured model magnitudes to their true magnitudes; the bias in the measurement is shown as a function of true magnitudes in Figure 6. There is appreciable scatter at a given magnitude, due both to the changing background and the different axis ratios. On average, however, the flux is underestimated by approximately 0.2 mag at $r = 12.5$ and <0.1 mag at $r = 15$ for simulated galaxies with an Sérsic index of 1. For an Sérsic index of 4, the flux is underestimated by as much as 0.3 mag at $r = 12.5$. The effect is more severe for simulated objects with an axis ratio of 1 than for an axis ratio of 0.5 (see Figure 6). The scale sizes of galaxies are similarly underestimated by as much as 20% for simulated galaxies with Sérsic index of 1, and 30% for an index of 4. Of course, the most massive elliptical or cD galaxies will have more extended envelopes, producing a larger effect than we have found here (Lauer et al. 2007).

6. IMPROVEMENTS IN PROCESSING OF SPECTROSCOPIC DATA

6.1. Correction of Instability in the Spectroscopic Flats

Spectroscopic flat fields for the blue camera in the first spectrograph contain an interference pattern produced by the dichroic. The thickness of the dichroic coating is believed to be sensitive to the ambient humidity, and moisture which enters the system during plate changes affects the instrument response, shifting the interference pattern in wavelength in unpredictable

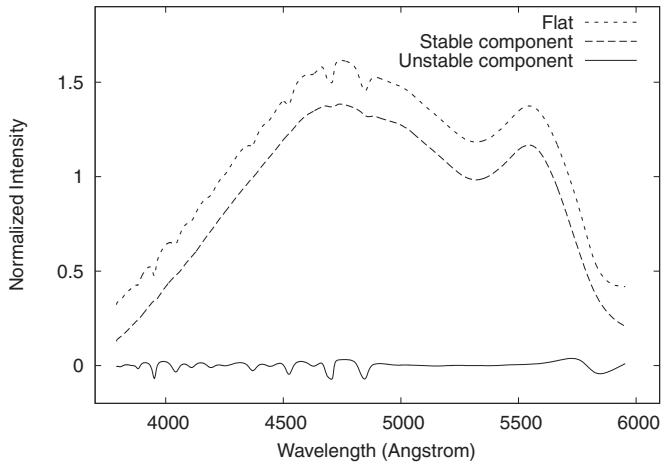

Figure 7. Decomposition of the flat field of the first blue spectrograph (upper curve) into stable (lower curve, offset slightly for clarity) and unstable (interference) components. The unstable component is close to zero, but shows wiggles at wavelengths that shift from one exposure to another.

ways on timescales comparable to the 900 s exposure time. The flats applied in processing were exposed several minutes prior to, or after, the science frames and therefore were not always representative of the true instrument response at the time of exposure. The interference pattern is most pronounced in the 3800–4100 Å region of the spectrum. If it shifts during an exposure, it will not be properly corrected by the flat field, causing significant distortion of blue absorption lines in stellar spectra, and systematically affecting estimates of metallicities and surface temperatures.

Flats obtained under different conditions were used to identify and model the stable and unstable (shifting) components of the flat, as shown in Figure 7. With this model in hand, we searched for shifts in the interference pattern over the typically 45 minute time a given plate was observed by comparing the results of the individual 15 minute exposures for each object. Thus, we took ratios of the extracted spectra from the separate exposures, and computed the median over all objects on a plate, giving results like those on the left-hand side of Figure 8. We fitted this ratio to the results expected from a shifting interference pattern (essentially a derivative of the shifting component in Figure 7), with the only free parameter being the amount of shift, and divided out this remaining component in each spectrum. The right-hand panel of Figure 8 shows that this technique removes the majority of the effects of the shifting interference. An example is shown in Figure 9, the spectrum of an A star observed on a plate where the interference term was particularly bad. The shapes of the absorption lines, especially Hϵ at 3970 Å, are much more regular in the new reductions.

6.2. Wavelength Calibration

The spectroscopic wavelength calibration is done quite accurately in SDSS, with typical errors of 2 km s^{-1} or better. As the DR6 paper describes, however, detailed analyses of stellar spectra revealed occasional errors that were substantially larger than this, especially in the blue end of the spectrum. The algorithms for fitting arc and sky lines were made more robust for DR6, and this improved the situation considerably. We have implemented two further improvements for DR7:

1. Spectroscopy is often done on nights with a moderate amount of moon. The bluest sky line used for wavelength calibration is an Hg line at 4046 Å, which is very close to a strong Fe I absorption line in the solar spectrum. Thus, when there is substantial moonlight in the sky spectrum, a fit to what is assumed to be an isolated emission line can be significantly biased, systematically skewing the wavelength solution at the blue end by as much as 20 km s^{-1}. In DR7, we now fitted this line to a linear combination of a Gaussian plus a stellar template including the absorption line, giving

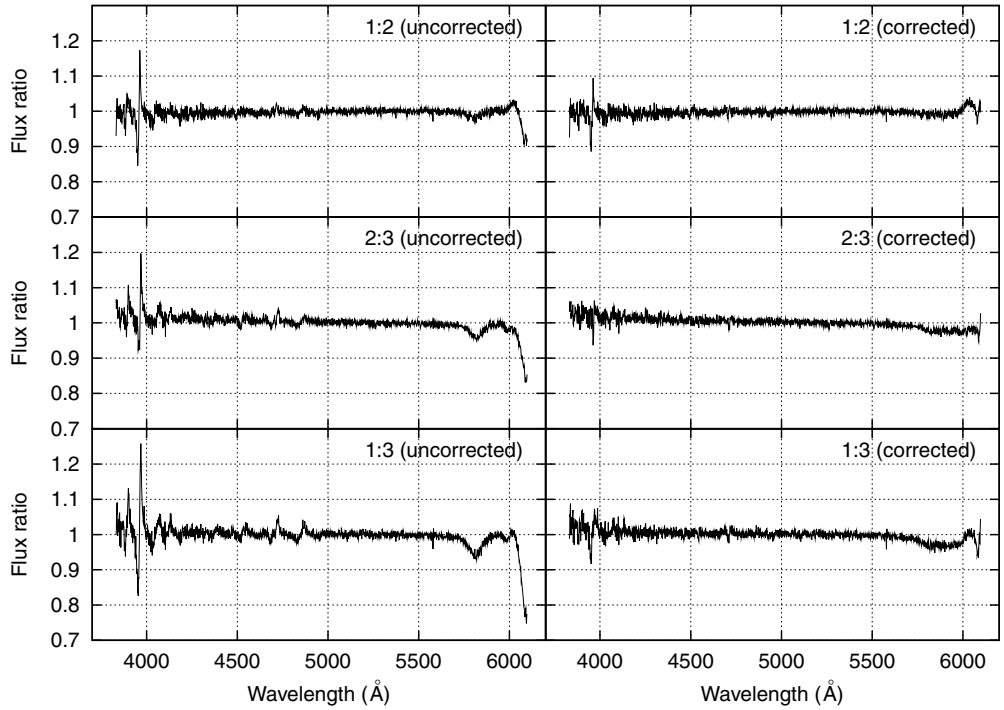

Figure 8. Median flux ratios over all objects in the three exposures of plate 1916, before (left) and after (right) correction for the moving interference filters. The ratio is fitted to the derivative of the interference component of the flat field (Figure 7) after allowing for an arbitrary wavelength shift.

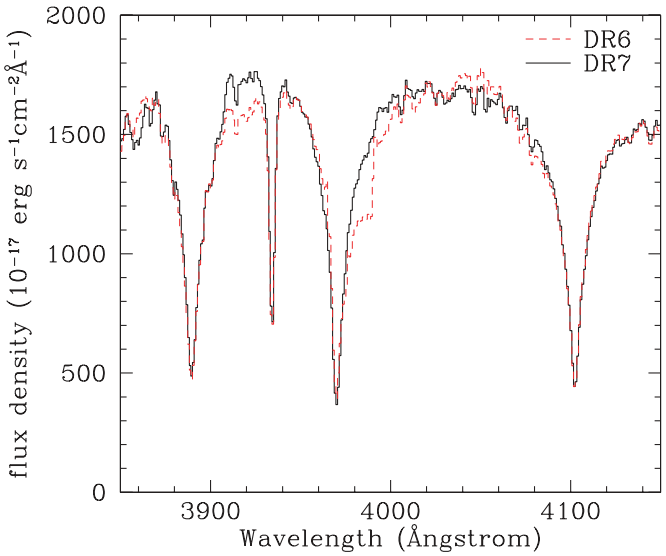

Figure 9. Spectrum of SDSS J172637.26+264127.6, an A0 star observed as part of SEGUE. The strong broad lines are due to Balmer absorption. The red dashed spectrum is that available in DR6, while the black solid spectrum is from DR7, with its improved flat field.

(A color version of this figure is available in the online journal.)

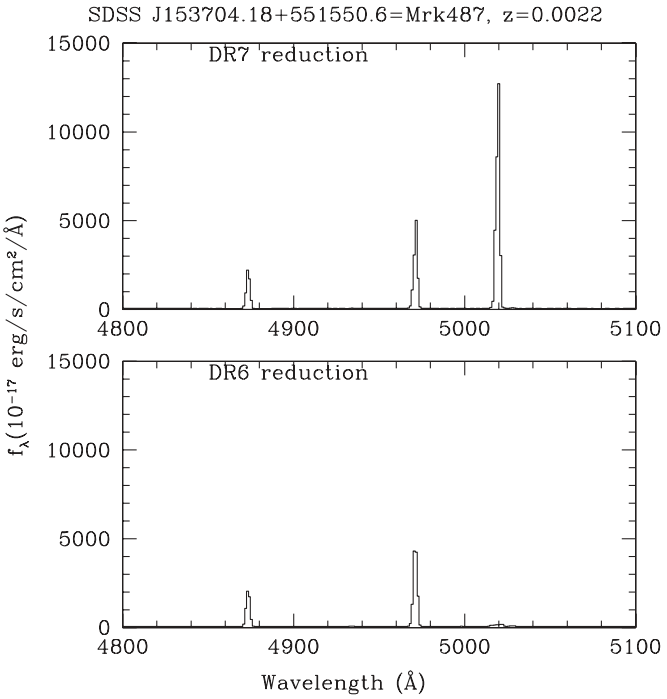

Figure 10. Spectra of the object SDSS J153704.18+551550.6 = Mrk487 in DR6 and DR7. The stronger [O III] emission line at 5020 Å was mistaken for a cosmic ray and clipped away completely in DR6, while the weaker line at 4970 Å was slightly affected. With the improved algorithm in DR7, the lines are not clipped.

an unbiased estimate of the wavelength of the line. In practice, bright moon affected 10 plates (listed in Yanny et al. 2009) out of a total of 410 SEGUE plates.

2. The sky and arc lines for each fiber are fitted to a wavelength solution; this is done independently for each fiber. This works well for the vast majority of plates. However, for a small fraction of plates, the arcs are weak (perhaps because the arc lamps themselves were faulty at that time, or because the petals which reflect the arc lamp light were not properly deployed), and the wavelength solution is poorly constrained. We therefore required that second- and higher-order terms in the wavelength solution be continuous functions of fiber number, to constrain the solution. We found that this produces much more robust wavelength solutions for those plates with weak arc observations, and has no substantial effect on the remaining plates.

The stellar spectral template library which gives the best radial velocity estimates is based on the ELODIE library (Prugniel & Soubiran 2001). We have removed one ELODIE template that gave velocities with a consistent offset from the rest of the library, as measured using the sample of ∼5000 stars with duplicate observations on each SEGUE plate pair. In order to provide more complete coverage in effective temperature, surface gravity, and metallicity for hot stars, we generated a grid of synthetic spectra using the models from Castelli & Kurucz (2003) over the same wavelength range and at the same resolving power as the spectra in the ELODIE library. This blue grid spans 6000–9500 K in 500 K increments, $-0.5 >$ [Fe/H] > -2.5 in increments of 0.5 dex, and $\log g$ of 2 and 4. We also added a grid of synthetic carbon-enhanced spectra (B. Plez 2008, private communication, using the stellar atmospheric code described by Gustafsson et al. 2008) at values of [Fe/H] between -1 and -4, [C/Fe] between 1 and 4, $\log g$ values between 2 and 4, and $T_{\rm eff}$ in the range 4000–6000 K. With these improvements, the radial velocity scatter in repeat observations for objects that match the carbon star templates is now the same as for the full sample.

The DR6 paper describes a 7 km s^{-1} systematic error in the radial velocities of stars (in the sense that the pipeline-reported velocities are too small). This is still with us in DR7; a correction is put into the outputs of the SEGUE Stellar Parameter Pipeline (SSPP; Lee et al. 2008a) but not elsewhere in the CAS or DAS. Beyond this problem, the plate-to-plate velocities of SEGUE stars have systematic errors of about 2 km s^{-1} in the mean. The rms velocity error of any given SEGUE star observation is about 5.5 km s^{-1} at $g = 18.5$, degrading to 12 km s^{-1} at $g = 19.5$.

6.3. Strong Unresolved Emission Lines

The spectroscopic pipeline combines observations of a given object on the red and blue spectrographs, and between the separate 15 minute exposures on the sky, by fitting a tightly constrained spline to the data, allowing discrepant points such as cosmic rays to be rejected. This spline requires as input the effective resolution of the spectra. As described in the DR6 paper, it did not do a perfect job; occasionally, very strong and sharp emission lines were erroneously rejected by this algorithm. This turned out to be due to the fact that the spline code did not adequately track the changing resolution of the spectra as a function of wavelength and fiber number. Including this effect significantly improved the behavior of this algorithm. Figure 10 shows an example spectrum of an object affected by this problem in DR6, and its improved counterpart in DR7, as is apparent by the correct 3:1 ratio of the 5007 Å and 4959 Å lines of [O III].

There is another problem, unfortunately not fixed in DR7, which has a similar effect. If the line is so bright that it is saturated in the individual 15 minute exposures of the spectrograph, it will also appear clipped. The flux value corresponding to saturation is a function of wavelength, but ranges from 2000 to 10,000 times 10^{-17} erg s^{-1} cm^{-2} Å^{-1} (the units in which spectral flux density in reported in the SDSS outputs). Unfortunately, such saturated pixels are not flagged as such, although usually

they are recognizable as having an inverse variance equal to zero. Luckily, objects with such strong emission lines are very rare, but the user should be aware of the possibility of objects with extremely strong emission lines and unphysical or unusual line ratios.

6.4. Improvements in the SEGUE Stellar Parameter Pipeline

There have been several improvements made in the SSPP (Lee et al. 2008a; 2008b; Allende Prieto et al. 2008a) since the release of DR6. In particular, in DR6, the SSPP underestimated metallicities (by about 0.3 dex) for stars approaching solar metallicity. This was fixed in DR7 by adding synthetic spectra with super-solar metallicities to two of the synthetic grid matching techniques (NGS1 and NGS2), and by recalibrating the CaIIK2, ACF, CaIIT, and ANNRR methods. See Table 5 in Lee et al. (2008a) for the naming convention for each technique. Two new techniques (ANNRR and CaIIK3) were also added to the SSPP metallicity estimation schemes, and contributed to the high-metallicity performance improvement.

Two methods, ACF and CaIIT, have been recalibrated to the "native" $g-r$ system, instead of making use of calibration on $B-V$, which required application of an uncertain transformation in color space. The ANNRR approach, which also tended to underestimate metallicity for near-solar metallicity stars, has been retrained on the SDSS/SEGUE spectra with improved stellar parameters, resulting in a better determination of the metallicity for metal-rich stars. Moreover, a neural network approach, based solely on noise-added synthetic spectra, has also been introduced. There remains a tendency for the SSPP to assign slightly higher metallicities for stars with [Fe/H] < -2.7. This offset is presently being calibrated out and will be corrected in SEGUE-2; see below. For more detailed descriptions of individual methods of the SSPP, we refer the interested reader to Lee et al. (2008a). Additionally, the pipeline now identifies cool main-sequence stars of low metallicity (late-K and M subdwarfs). The stars are assigned metallicity classes and spectral subtypes following the classification system of Lépine et al. (2007). Cool and ultracool subdwarfs are classified as subdwarfs (sdK, sdM), extreme subdwarfs (esdK, esdM), and ultrasubdwarfs (usdK, usdM) in order of decreasing metal content. The classification is based on the absolute and relative values of the TiO and CaH molecular band strengths, and derived from fits to K–M dwarf and K–M subdwarf spectral templates.

A number of open and globular clusters have been observed photometrically and spectroscopically with the SDSS instruments to evaluate the performance of the SSPP (Lee et al. 2008b). In addition, high-resolution spectra have been obtained for about 100 field stars included in the SDSS, and used to expand the SSPP checks over a wider parameter space (Allende Prieto et al. 2008a).

7. LOOKING AHEAD TO SDSS-III

This paper marks the release of the final data of SDSS-II. The original SDSS science goals (York et al. 2000) included five-band imaging over 10^4 deg^2 with 2% rms errors or better in photometric calibration, and spectroscopy of 10^6 galaxies and 10^5 quasars. We have met these goals, and have in addition carried out extensive stellar spectroscopy of close to half a million stars, and repeat imaging over 250 deg^2 to search for supernovae. Over 2200 refereed papers have been published to date using SDSS data or results, on subjects ranging from the large-scale distribution of galaxies to distant quasars to substructure in the Galactic halo to surveys of white dwarfs to the color distribution of main belt asteroids.

The SDSS telescope has started a new operational phase, called SDSS-III, which will include four surveys with the 2.5 m telescope through 2014:

1. SEGUE-2 extends the science goals of SEGUE with the same instrumentation and data processing pipelines, but targets fainter stars to study the distant halo. It will increase the number of distant halo stars by a factor of 2.5 with respect to the results of SDSS and SDSS-II.
2. The Baryon Oscillation Spectroscopic Survey (BOSS) will perform spectroscopy of 1.5 million luminous red galaxies to $z \approx 0.7$ and 160,000 quasars with $2.3 < z < 3$ to measure the scale of the baryon oscillation signal in the correlation function as a function of redshift (Schlegel et al. 2007).
3. The Multi-object APO Radial Velocity Exoplanet Large-area Survey (MARVELS) will monitor the radial velocities of 11,000 bright stars to search for the signature of planets with periods ranging from several hours to two years (Ge et al. 2008).
4. The APO Galactic Evolution Experiment (APOGEE) will perform $R \approx 20,000$ H-band spectroscopy of 10^5 giant stars to $H = 13.5$ for detailed radial velocity and chemical studies of the Milky Way (Majewski et al. 2008; Allende Prieto et al. 2008b).

These data will be made public in a series of data releases, following the pattern established by SDSS and SDSS-II.

This paper represents the end of SDSS-II, the culmination of a project taking two decades and involving an enormous number of scientists from all over the world. We dedicate this paper to colleagues who made essential contributions to the SDSS but are no longer with us: John N. Bahcall, Don Baldwin, Norm Cole, Arthur Davidsen, Jim Gray, Bohdan Paczyński, and David N. Schramm. The successful completion of this project is in large part a reflection of the hard work and intellectual capital they put into it.

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

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The participating institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

REFERENCES

Abazajian, K., et al. 2003, AJ, 126, 2018 (DR1 paper)
Abazajian, K., et al. 2004, AJ, 128, 502 (DR2 paper)
Abazajian, K., et al. 2005, AJ, 129, 1755 (DR3 paper)
Adelman-McCarthy, J. K., et al. 2006, ApJS, 162, 38 (DR4 paper)
Adelman-McCarthy, J. K., et al. 2007, ApJS, 172, 634 (DR5 paper)
Adelman-McCarthy, J. K., et al. 2008, ApJS, 175, 297 (DR6 paper)
Allende Prieto, C., et al. 2008a, AJ, 136, 2070
Allende Prieto, C., et al. 2008b, AN, 329, 1018
An, D., et al. 2008, ApJS, 179, 326
Anderson, S. F., et al. 2007, AJ, 133, 313
Becker, R. H., White, R. L., & Helfand, D. J. 1995, ApJ, 450, 559
Bernardi, M., et al. 2007, AJ, 133, 1954
Blanton, M. R., Lin, H., Lupton, R. H., Maley, F. M., Young, N., Zehavi, I., & Loveday, J. 2003, AJ, 125, 2276
Blanton, M. R., et al. 2005, AJ, 129, 2562
Bramich, D. M., et al. 2008, MNRAS, 386, 887
Budavári, T., Szalay, A. S., Connolly, A. J., Csabai, I., & Dickinson, M. 2000, AJ, 120, 1588
Castelli, F., & Kurucz, R. L. 2003, in IAU Symp. 210, ed. N. Piskunov, W. W. Weiss, & D. F. Gray (San Francisco, CA: ASP), 20
Csabai, I., Dobos, L., Trencséni, M., Herczegh, G., Józsa, P., Purger, N., Budavári, T., & Szalay, A. S. 2007, Astron. Nachr., 328, 852
Csabai, I., et al. 2003, AJ, 125, 580
Cunha, C. E., Lima, M., Oyaizu, H., Frieman, J., & Lin, H. 2008, arXiv:0810.2991
Davis, M., et al. 2007, ApJ, 660, L1
Eisenstein, D. J., et al. 2001, AJ, 122, 2267
Finkbeiner, D. P., et al. 2004, AJ, 128, 2577
Frieman, J. A., et al. 2008, AJ, 135, 338
Fukugita, M., Ichikawa, T., Gunn, J. E., Doi, M., Shimasaku, K., & Schneider, D. P. 1996, AJ, 111, 1748
Ge, J., et al. 2008, in AAS DPS meeting, 40, 1.06
Gunn, J. E., et al. 1998, AJ, 116, 3040
Gunn, J. E., et al. 2006, AJ, 131, 2332
Gustafsson, B., Edvardsson, B., Eriksson, K., Jørgensen, U. G., Nordlund, Å., & Plez, B. 2008, A&A, 486, 951
Holtzman, J. A., et al. 2008, AJ, 136, 2306
Høg, E., et al. 2000, A&A, 355, L27
Hogg, D. W., Finkbeiner, D. P., Schlegel, D. J., & Gunn, J. E. 2001, AJ, 122, 2129
Ivezić, Ž., et al. 2004, Astron. Nachr., 325, 583
Ivezić, Ž., et al. 2007, AJ, 134, 973
Jiang, L., et al. 2008, AJ, 135, 1057
Kaiser, N., et al. 2002, SPIE, 4836, 154
Lauer, T. R., et al. 2007, ApJ, 662, 808
Lee, Y. S., et al. 2008a, AJ, 136, 2022
Lee, Y. S., et al. 2008b, AJ, 136, 2050
Lépine, S., Rich, R. M., & Shara, M. M. 2007, ApJ, 669, 1235
Lilly, S., et al. 2007, ApJS, 172, 70
Lisker, T., Grebel, E. K., Binggeli, B., & Glatt, K. 2007, ApJ, 660, 1186
Lupton, R. H., Gunn, J. E., Ivezić, Ž., Knapp, G. R., Kent, S., & Yasuda, N. 2001, in ASP Conf. Proc. 238, Astronomical Data Analysis Software and Systems X, ed. F. R. Harnden, Jr., F. A. Primini, & H. E. Payne (San Francisco, CA: ASP), 269
Lupton, R. H., Gunn, J. E., & Szalay, A. S. 1999, AJ, 118, 1406
Magnier, E. 2006, in The Advanced Maui Optical and Space Surveillance Technologies Conference, ed. S. Ryan (The Maui Economic Development Board), E50
Majewski, S. R., et al. 2008, BAAS, 39, 962
Mandelbaum, R., et al. 2005, MNRAS, 361, 1287
Mandelbaum, R., et al. 2008, MNRAS, 386, 781
Masjedi, M., et al. 2006, ApJ, 644, 54
McGehee, P. M., West, A. A., Smith, J. A., Anderson, K. S. J., & Brinkmann, J. 2005, AJ, 130, 1752
Munn, J. A., et al. 2004, AJ, 127, 3034
Munn, J. A., et al. 2007, AJ, 136, 895
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Oyaizu, H., Lima, M., Cunha, C. E., Lin, H., Frieman, J., & Sheldon, E. S. 2008, ApJ, 674, 7680
Padmanabhan, N., et al. 2008, ApJ, 674, 1217
Petrosian, V. 1976, ApJ, 210, L53
Pier, J. R., Munn, J. A., Hindsley, R. B., Hennessy, G. S., Kent, S. M., Lupton, R. H., & Ivezić, Ž. 2003, AJ, 125, 1559
Prugniel, P., & Soubiran, C. 2001, A&A, 369, 1048
Richards, G. T., et al. 2002, AJ, 123, 2945
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525 (SFD)
Schlegel, D. J., et al. 2007, BAAS, 39, 966
Smith, J. A., et al. 2002, AJ, 123, 2121
Stetson, P. B. 1987, PASP, 99, 191
Stetson, P. B. 1994, PASP, 106, 250
Stoughton, C., et al. 2002, AJ, 123, 485 (EDR paper)
Strauss, M. A., et al. 2002, AJ, 124, 1810
Subbarao, M., Frieman, J., Bernardi, M., Loveday, J., Nichol, B., Castander, F., & Meiksin, A. 2002, SPIE, 4847, 452
Thakar, A. R., Szalay, A., Fekete, G., & Gray, J. 2008, Comput. Sci. Eng., 10, 30
Tucker, D., et al. 2006, AN, 327, 8212
York, D. G., et al. 2000, AJ, 120, 1579
Yanny, B., et al. 2009, arXiv:0902.1781
Zacharias, N., Urban, S. E., Zacharias, M. I., Wycoff, G. L., Hall, D. M., Monet, D. G., & Raffert, T. J. 2004, AJ, 127, 3043
Zacharias, N., et al. 2000, AJ, 120, 2131