

The Sloan Digital Sky Survey(s)

Before we begin

- Any questions about the homework?
- Seminar topics
- Next two weeks (classes & travel)

Next two weeks

- I'm away this Friday, but here tomorrow
- Monday class will be on WISE
- I'm then away Tuesday through Friday next week
- Wednesday class will be on PHAT (by Ben Williams)
- No class on Monday, Feb 16th
- Wednesday, Feb 18th will be the first two seminars
 - Any volunteers?

Seminars

- CFHT-LS
- Kepler
- Fermi
- UKIDSS
- GALEX (Joachim)
- CRTS
- PTF (Paul)
- VLA FIRST
- 2MASS
- RAVE
- VISTA (Bryce)
- MWA (Patti)

SDSS I, II, III and IV

- A series of optical photometric and spectroscopic surveys using a dedicated 2.5m telescope at the Apache Point Observatory, co-funded by the Sloan Foundation and the NSF. Started in late 1990-ies and still ongoing today.
- Conducted by the SDSS Collaboration. UW was a member from the beginning.

Overview

- The telescope and instruments
- The software and data
- The SDSS surveys
- SDSS science legacy

Historical Context

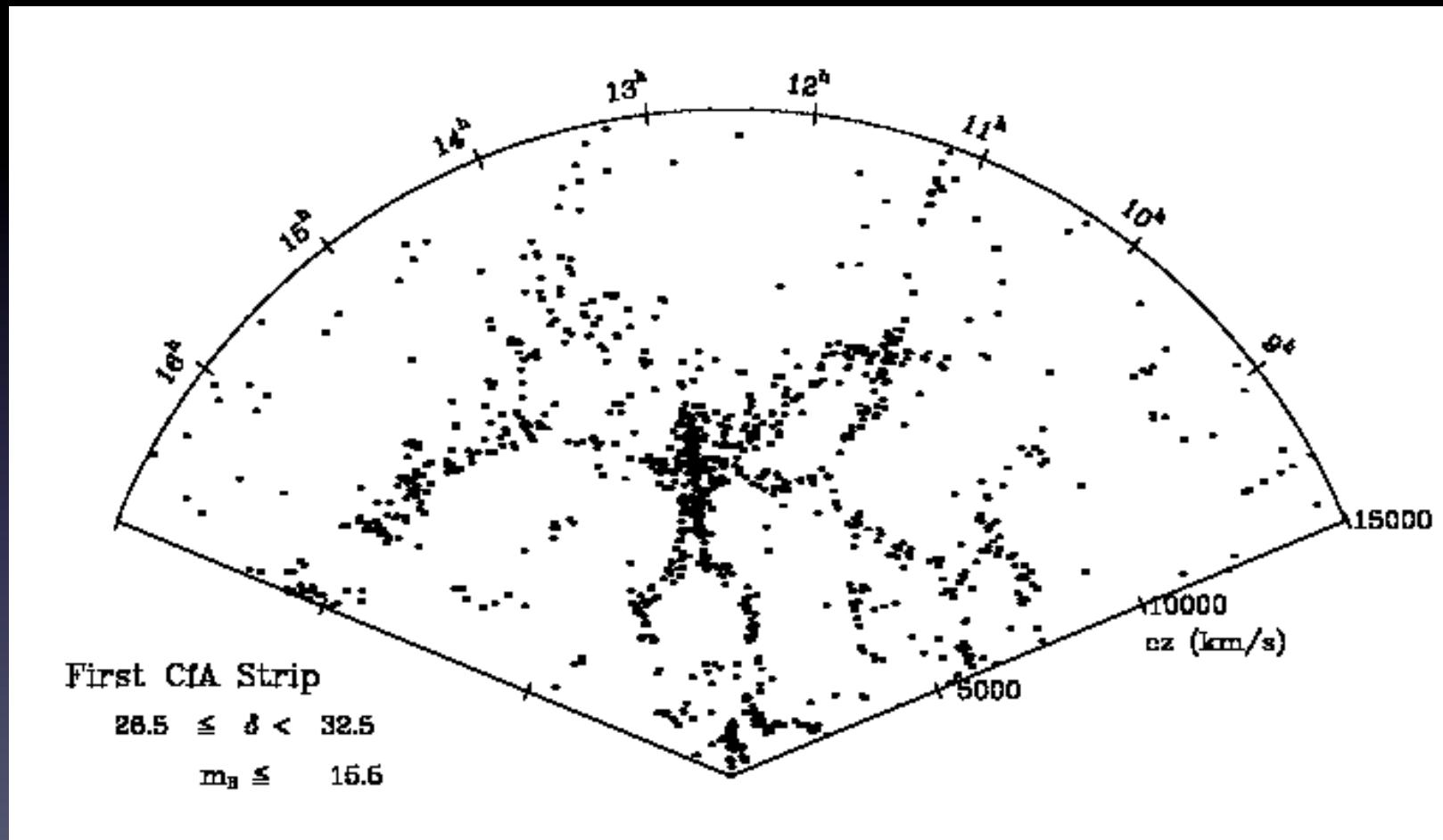


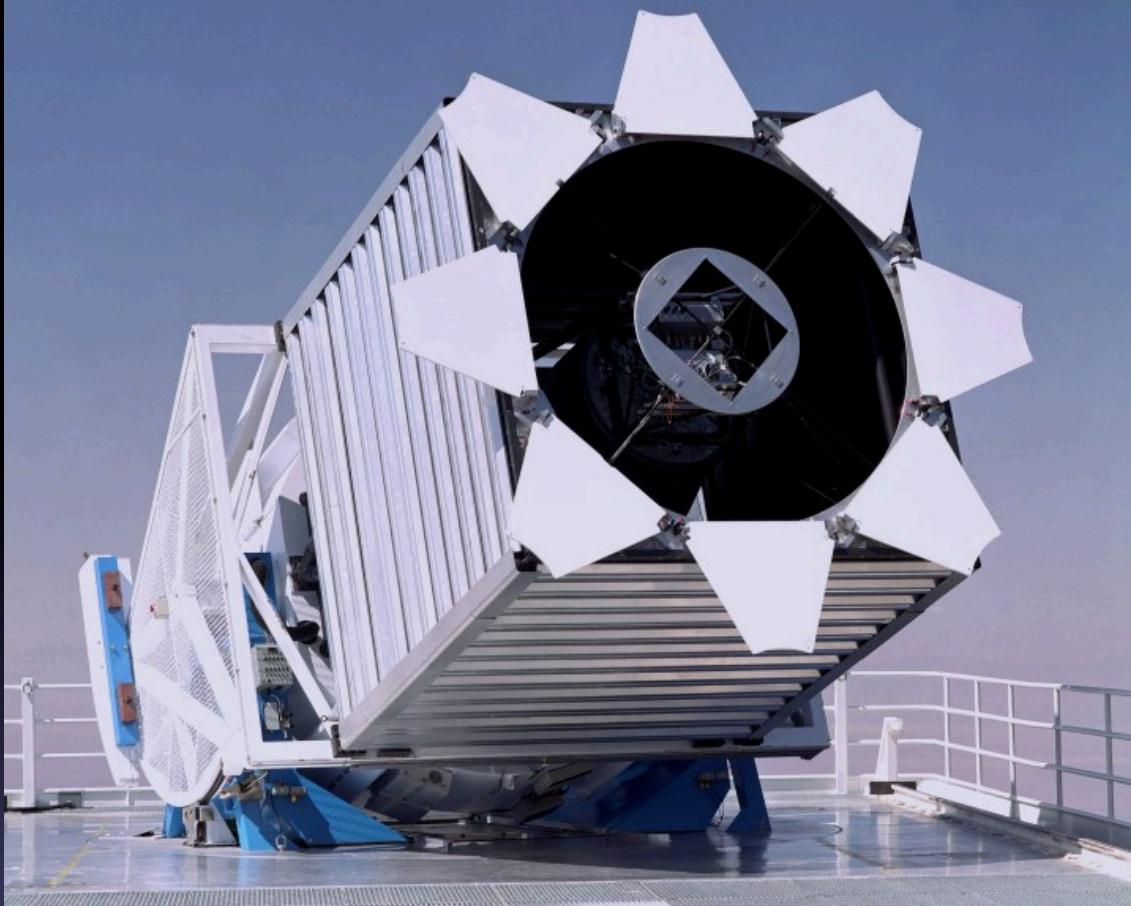
Image credit: Smithsonian Astrophysical Observatory
Geller and Huchra 1989, Science 246, 897

How it all got began

"In 1988, a team of astrophysicists gathered together for the task of designing a next-generation redshift survey – one that would target both galaxies and quasars. In order to achieve the highest level of homogeneity in these two redshift samples, it was concluded that a dedicated imaging survey would be needed from which target galaxies and quasars would be selected, and that imaging and spectroscopy could be done with the same telescope switching between the two observing modes. Substantial improvement beyond existing surveys dictated an increase by a factor of 100 in terms of the number of targets available at the time – in other words, a survey of 1 million galaxy redshifts.

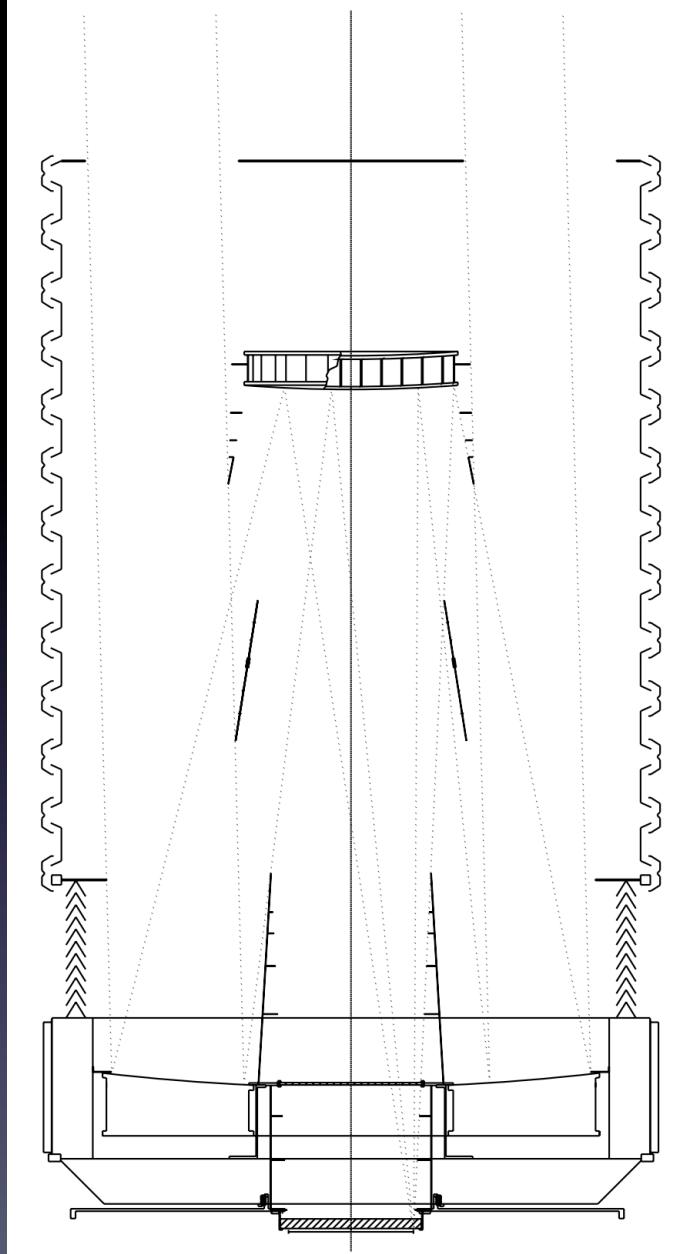
This survey, the Sloan Digital Sky Survey (SDSS; York et al. 2000), is now underway, having begun standard operations in 2000 April, and is planned to last 5 years."

- New telescope
- A camera
- Spectrograph
- Data processing system
- People

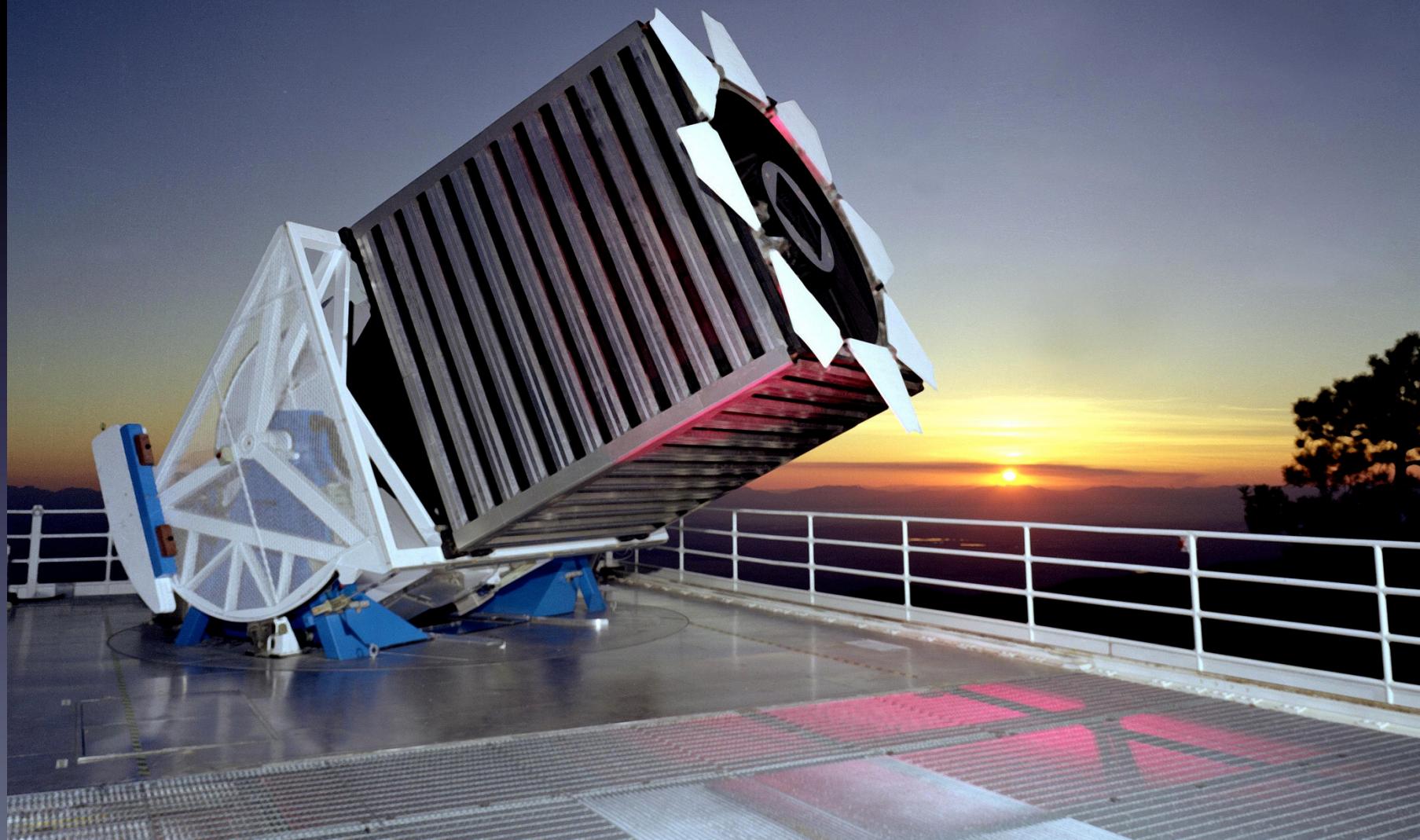


A dedicated 2.5 m f/5 modified Ritchey-Chretien altitude-azimuth telescope.

A 1.08 m secondary mirror and two corrector lenses result in a 3° distortion-free field of view.



Gunn et al. (2006): "The 2.5 m Telescope of the Sloan Digital Sky Survey"

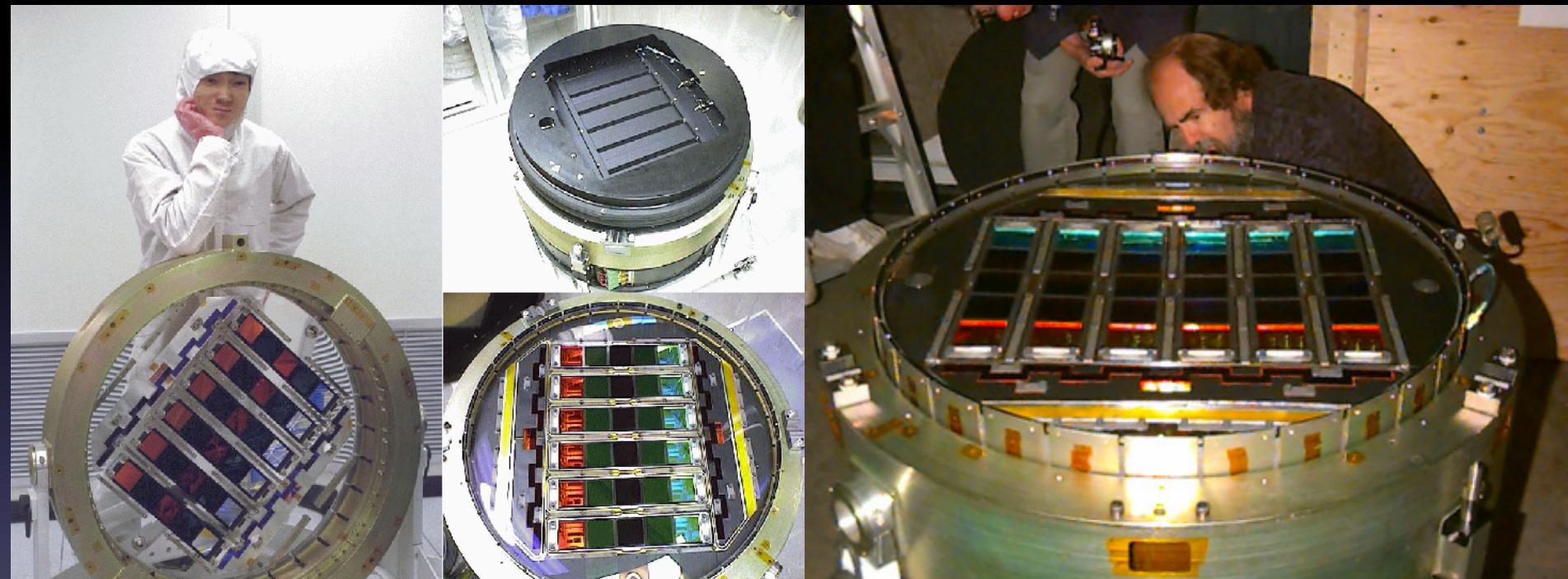




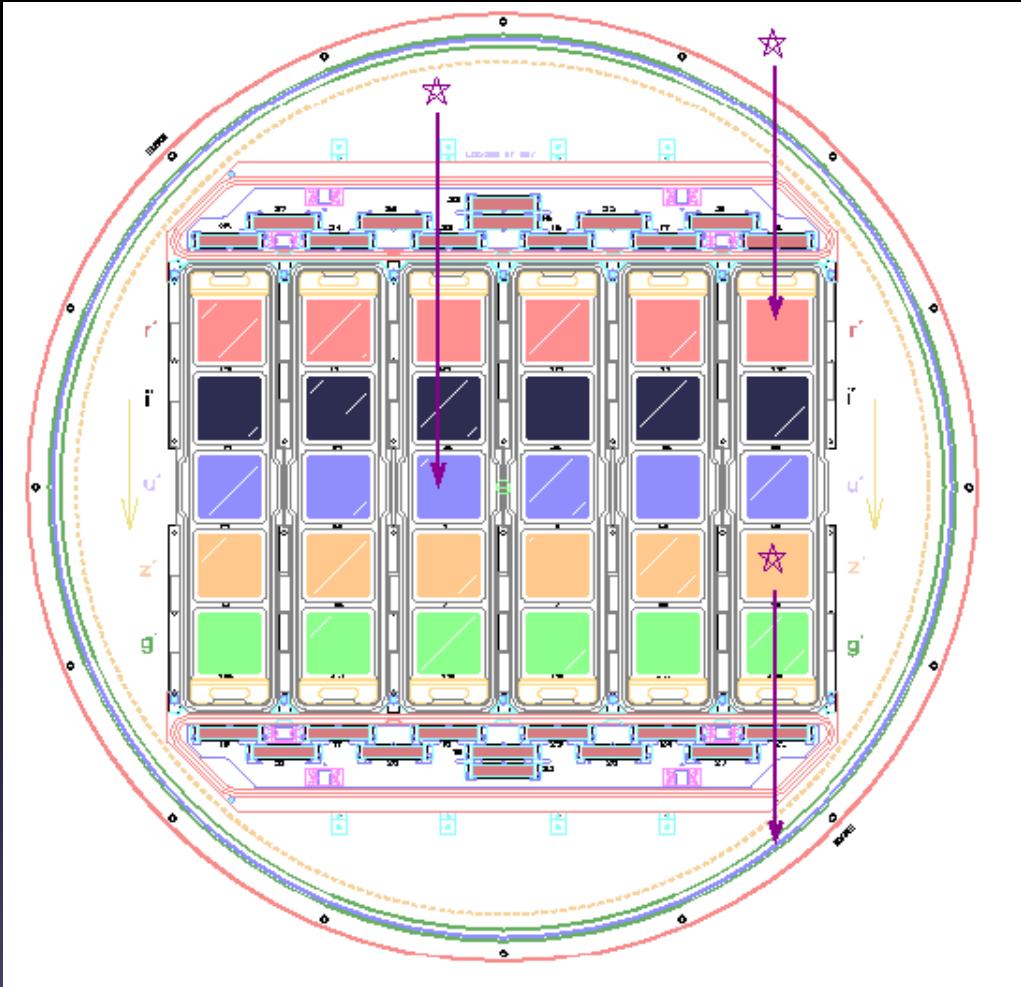
Located at Apache Point Observatory in south east New Mexico. Total seeing ~1.4".

*Latitude 32° 46' 49.30" N
Longitude 105° 49' 13.50" W
Elevation 2788m*

SDSS Camera



30 2k x 2x sensors in a 6 x 5 array, with five independent filters.
22 smaller chips for astrometric calibration
2 chips for focus monitoring



6 “camera columns”, with five rows of CCDs.
Each row has the same filter placed over the CCD (more later).

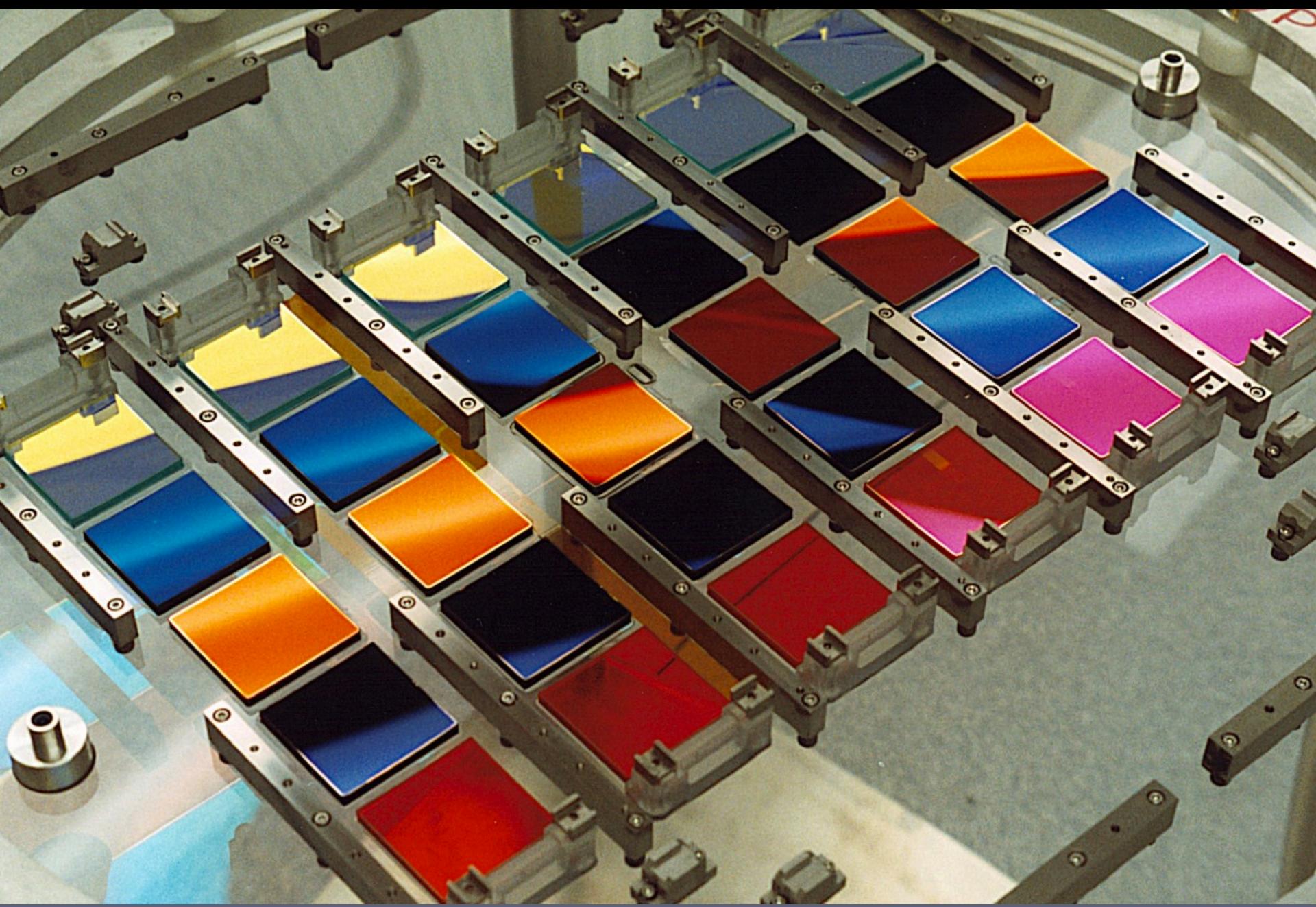
Operating temperature of -60 degC

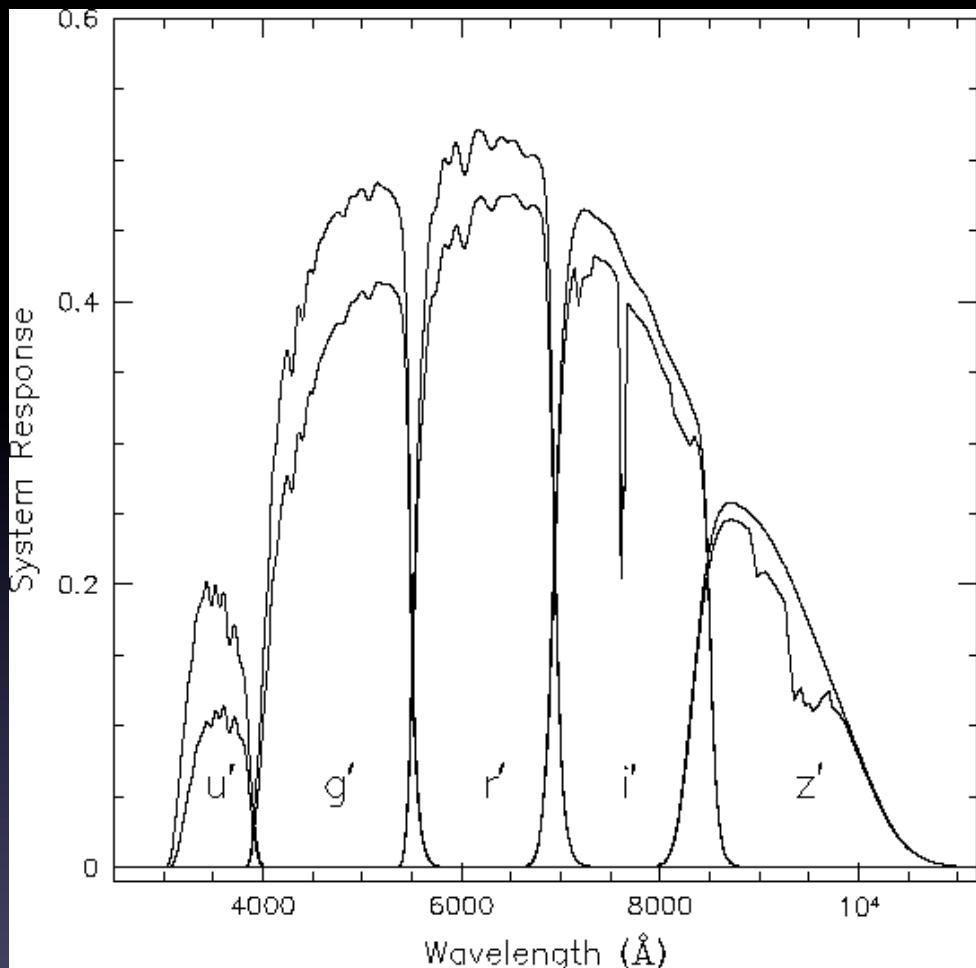
The SDSS imaging data are taken in time-delay-and-integrate (TDI) mode at the sidereal rate almost simultaneously in five bands.

The sky tracks through 5 CCD detectors in succession, each located behind a different filter.

Exposure time is 53.9 seconds (for each CCD), and the

See “TDI” in <http://www.sdss.org/dr12/help/glossary/>





Gunn-Thuan ugriz photometric filters.

Similar to UBVRI but with stronger rejection of atmospheric lines (O I and Hg I)

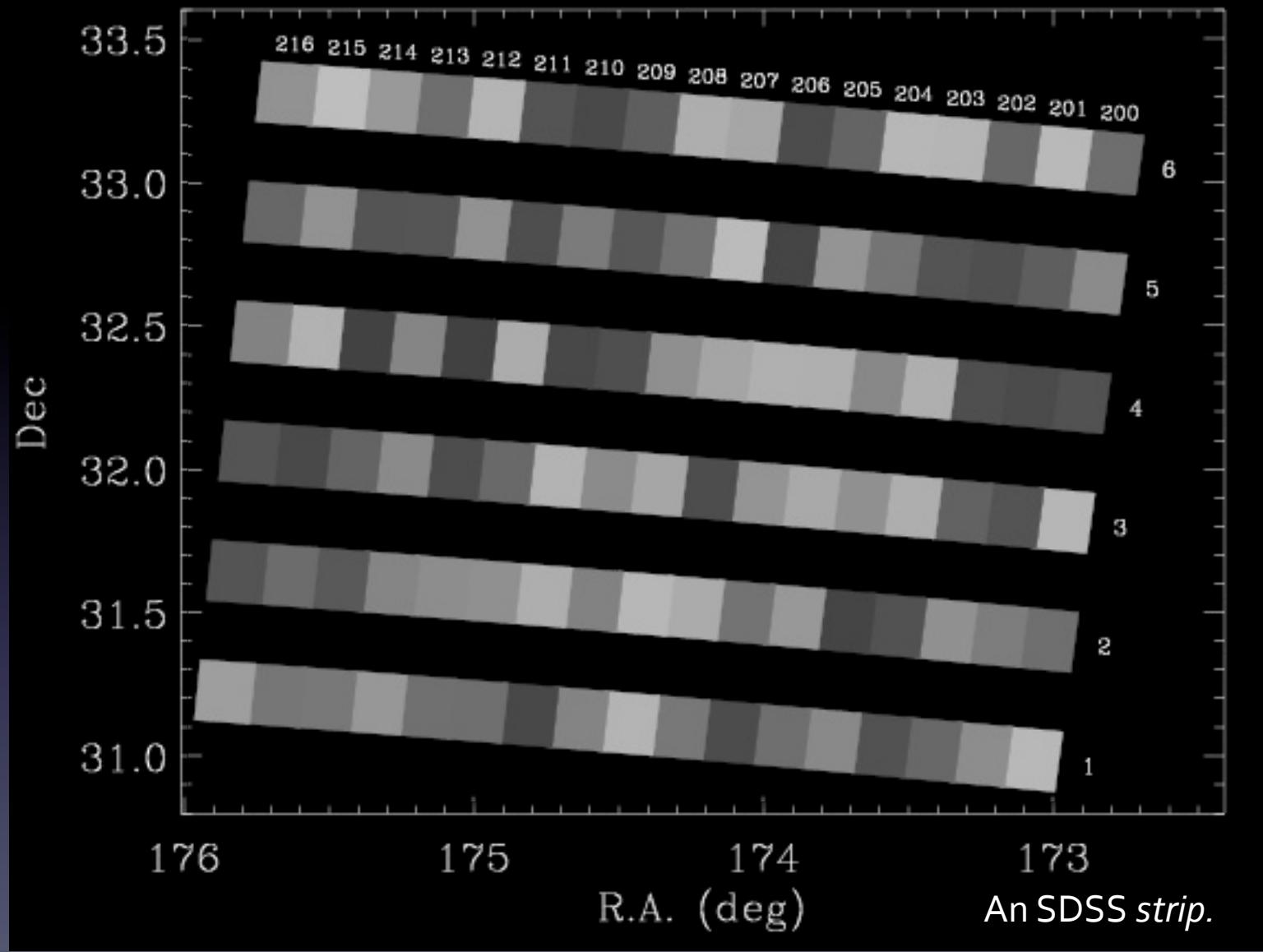
AB photometric system (Oke & Gunn, 1983).

Fukigita et al. (1996): "The Sloan Digital Sky Survey Photometric System"

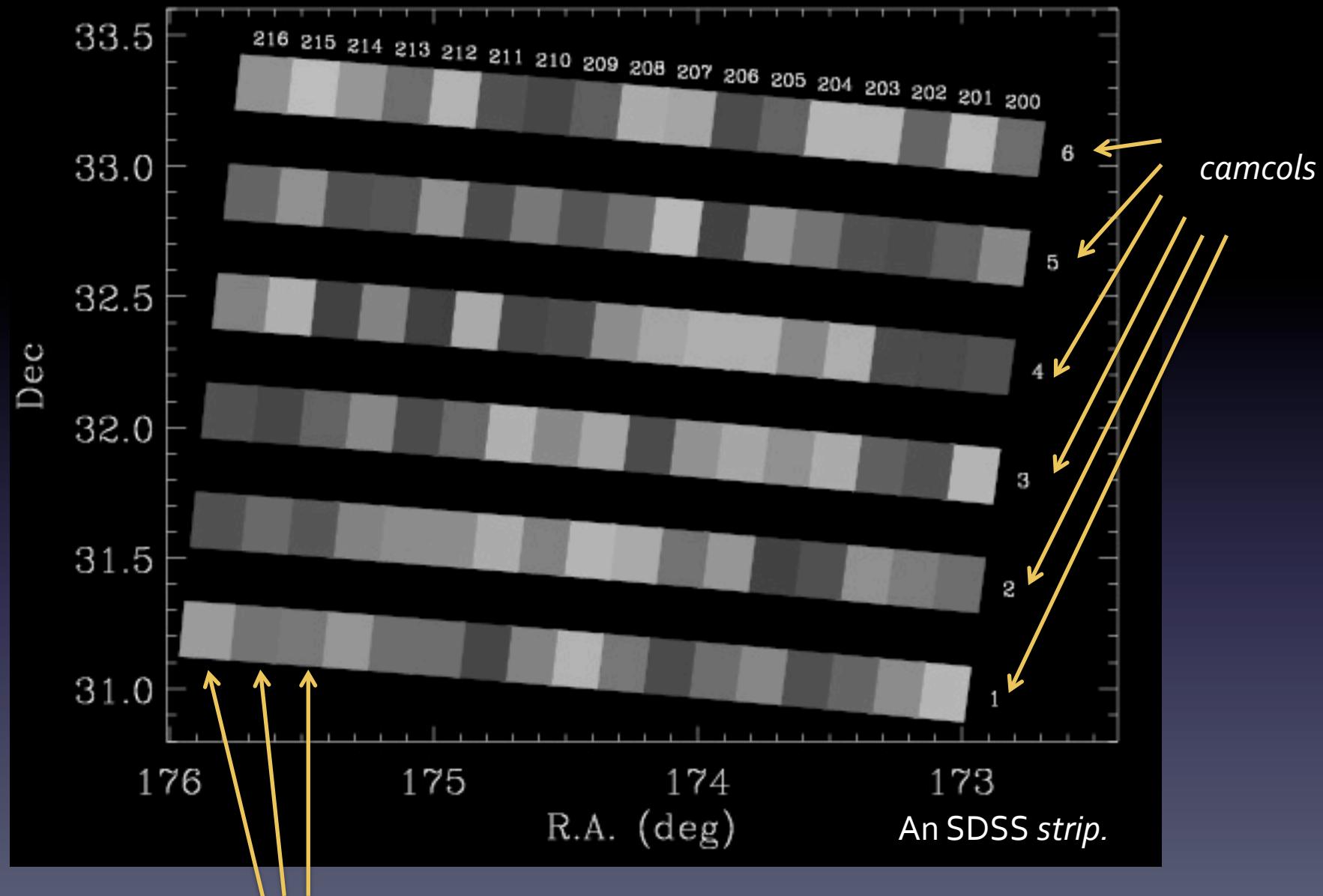
<i>u</i>	<i>g</i>	<i>r</i>	<i>i</i>	<i>z</i>
3551Å	4686Å	6165Å	7481Å	8931Å
22.0	22.2	22.2	21.3	20.5

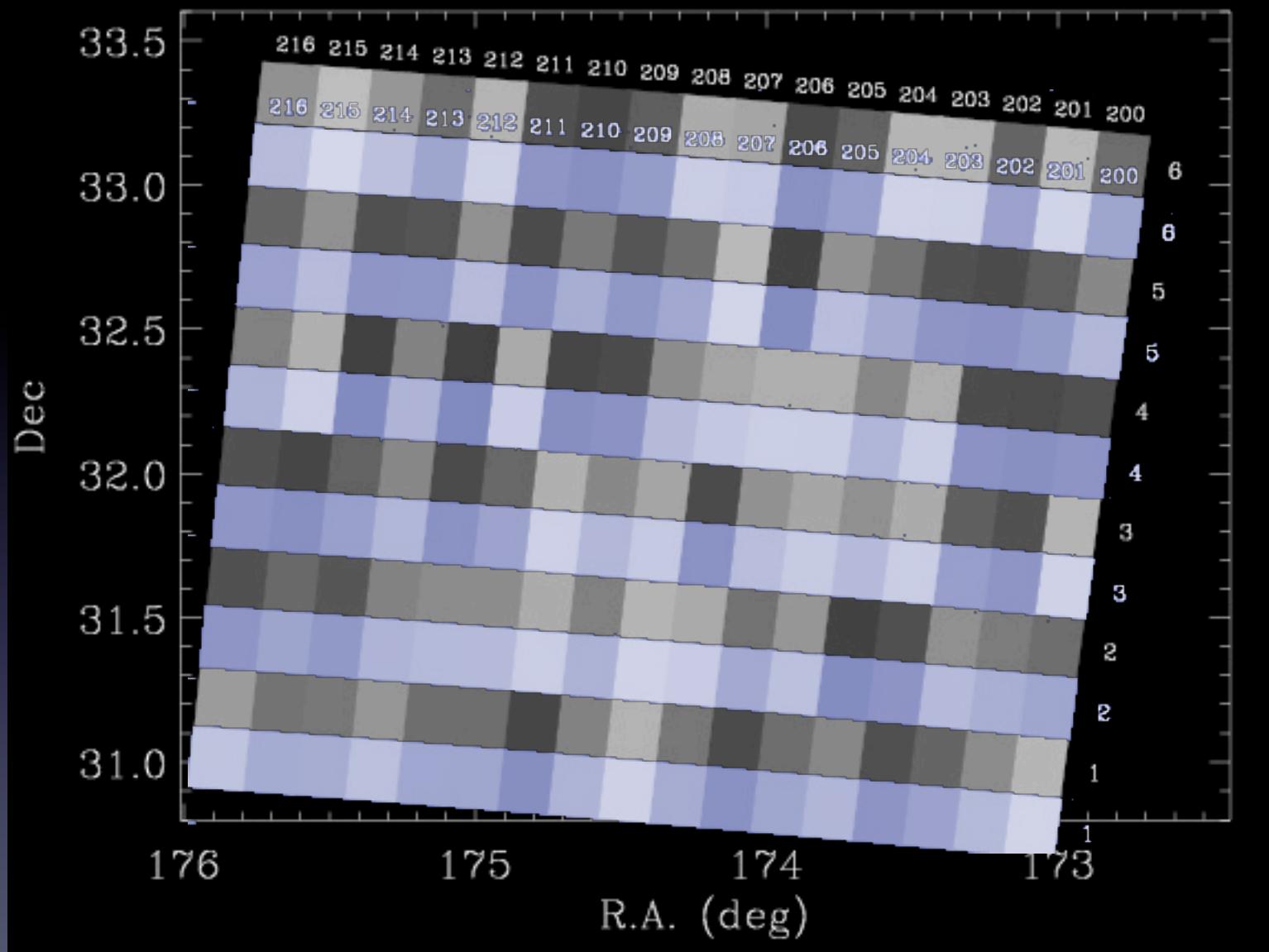
Average wavelengths and magnitude limits (95% detection repeatability for point sources)

Observing pattern in one *observing run*.

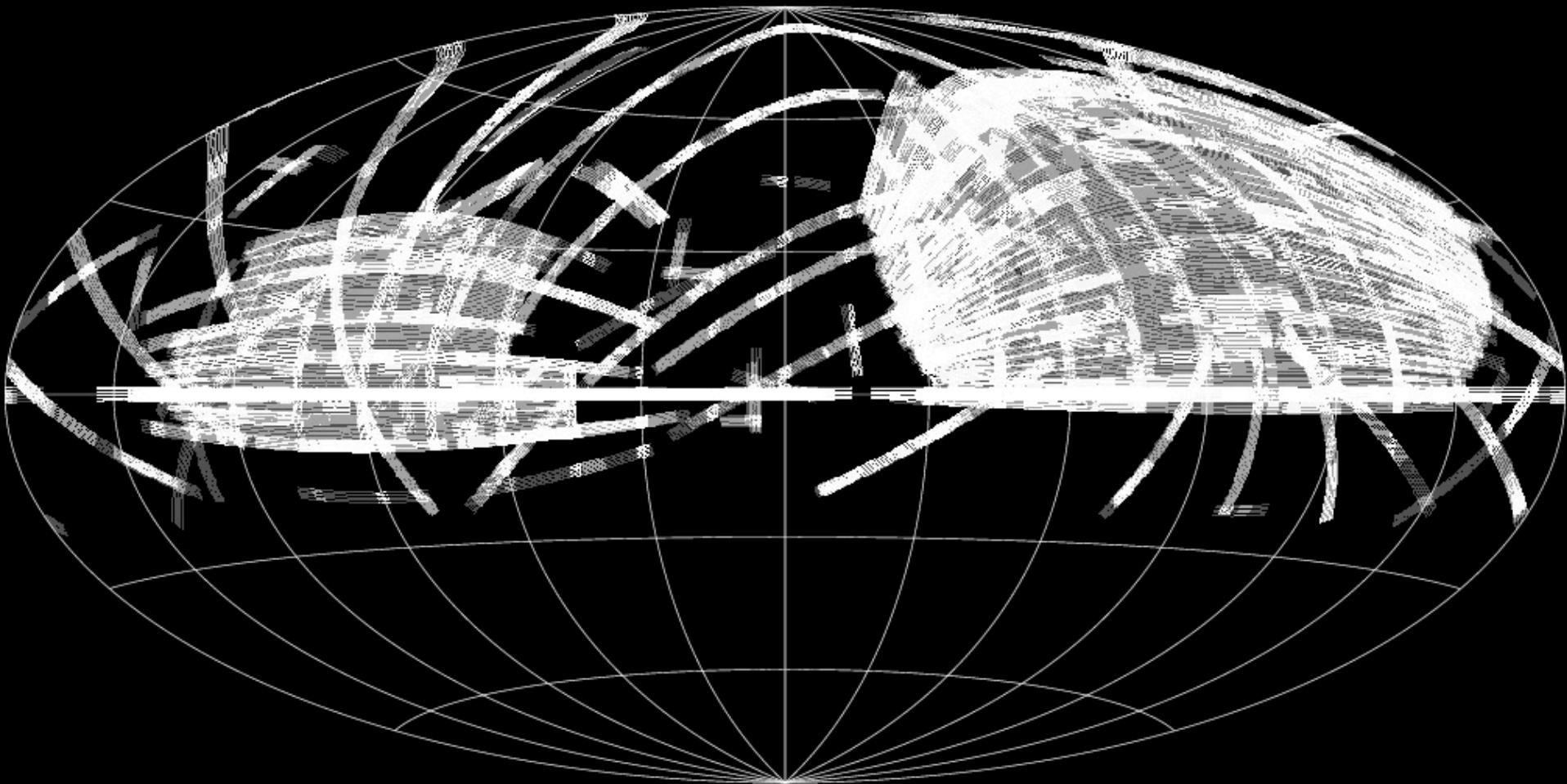


Observing pattern in one *observing run*.





Two *strips*, observed with a slight offset, make a single *stripe*.



One SDSS field



2048 px ($\approx 13'$)

frame-r-004570-4-0135

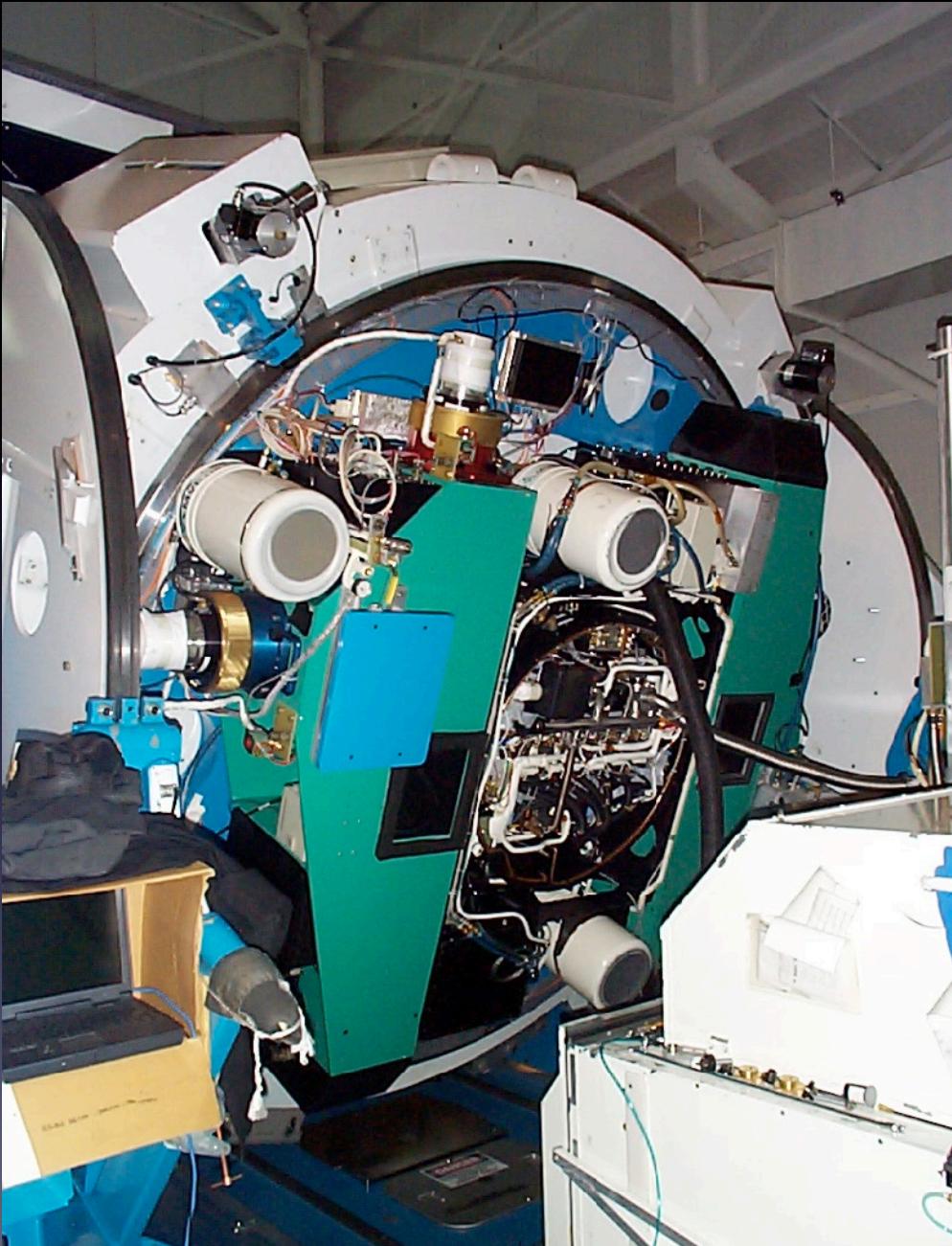
1489 px
($\approx 10'$)

To Summarize: How SDSS Images The Sky

- The SDSS imaging camera scanned the sky in strips along great circles. Each strip consists of six parallel scanlines, 13 arcmin wide, with gaps of about the same width. Thus two strips, offset slightly from each other, together make a single stripe 2.5 degrees wide.
- Each scanline includes data in all five filters, ugriz. The fundamental units of SDSS images are fields into which the scanlines are divided (with some overlap). Each is 10 by 13 arcminutes, corresponding to 2048 by 1489 pixels.
- Each field can be uniquely identified by a sequence of three numbers:
 - the *run number*, which identifies the specific scan,
 - the *camera column*, or “camcol,” a number from 1 to 6, identifying the scanline within the run, and
 - the *field number*. The field number typically starts at 11 (after an initial rampup time), and can be as large as 800 for particularly long runs.

Spectroscopy

- SDSS is also (very prolific) spectroscopy machine. The camera is periodically taken off the telescope to use the spectrographs.
- The original survey begun with two identical multi-object spectrographs, allowing it to take spectra of hundreds of objects at a time. They were upgraded in the SDSS III era (and additional spectrographs were added as well).
- Multi-object functionality is accomplished by placing optical fibers on the location in the focal plane where the objects' light gets focused. The light travels over the fibers into the spectrograph for dispersion and imaging of spectra.

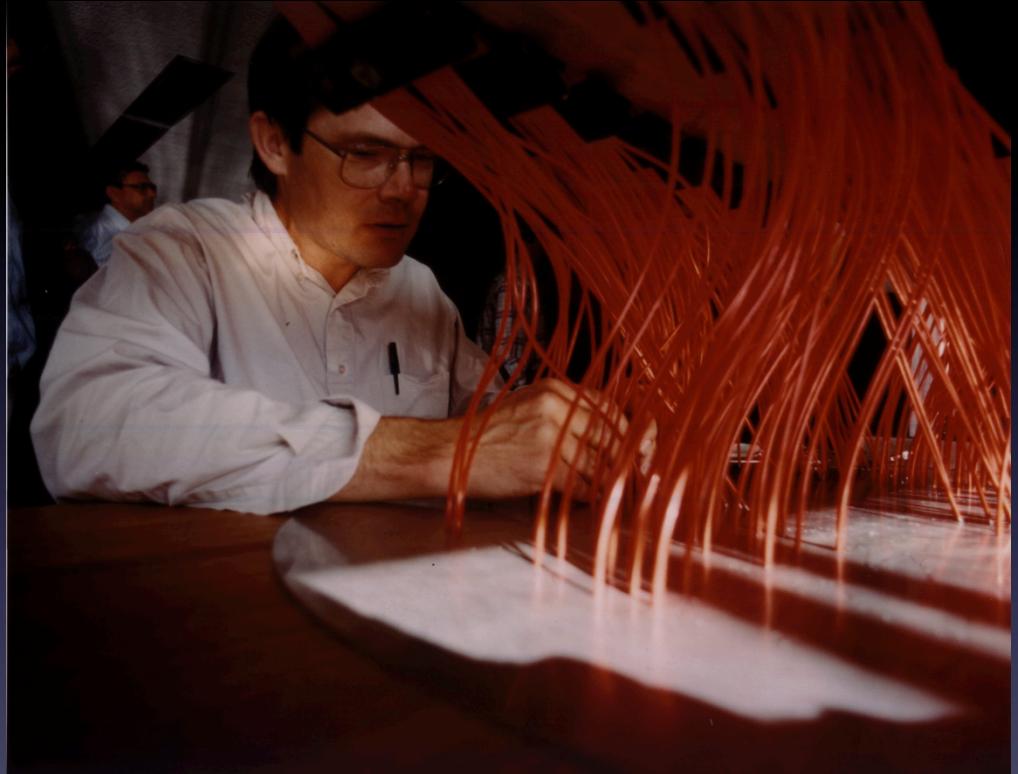


Two SDSS spectrograph are visible on the back side of the telescope (green boxes).

The black ring in the middle is the back of the camera. When spectra are being taken, the camera is removed and replaced with a cartridge that holds the optical fibers (below).



Positioning the Fibers: Plug Plates!



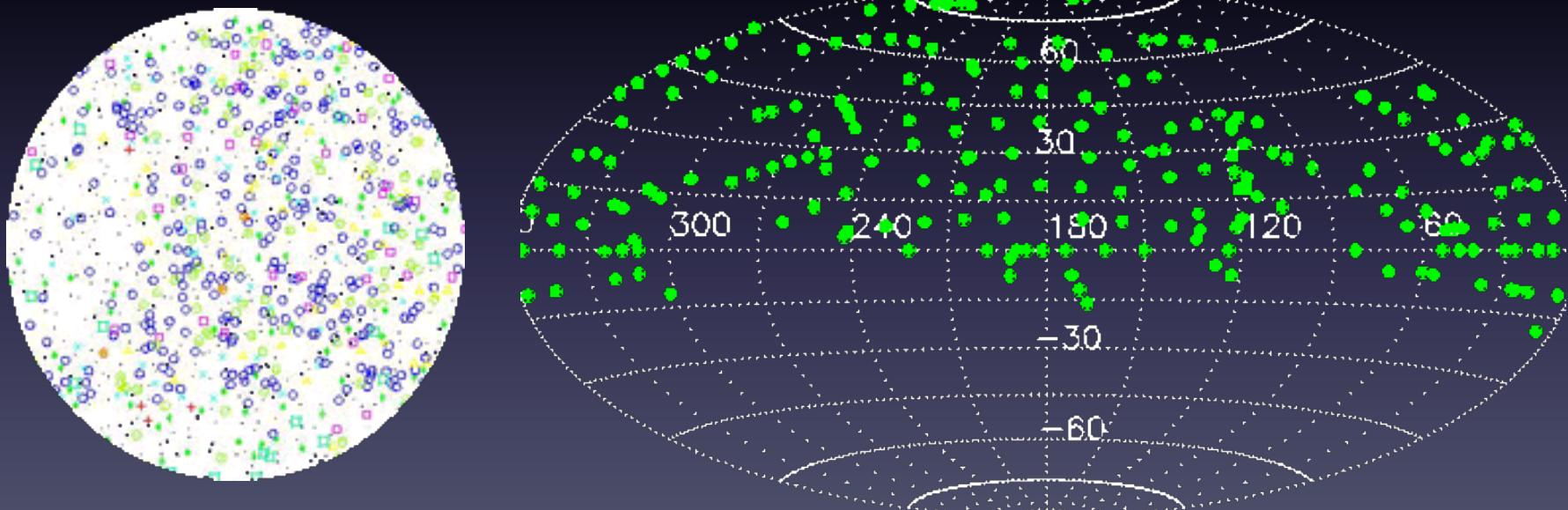
Making plug-plates at the UW machine shop:

<http://goo.gl/RPoAeS>

Useful beyond astronomy!



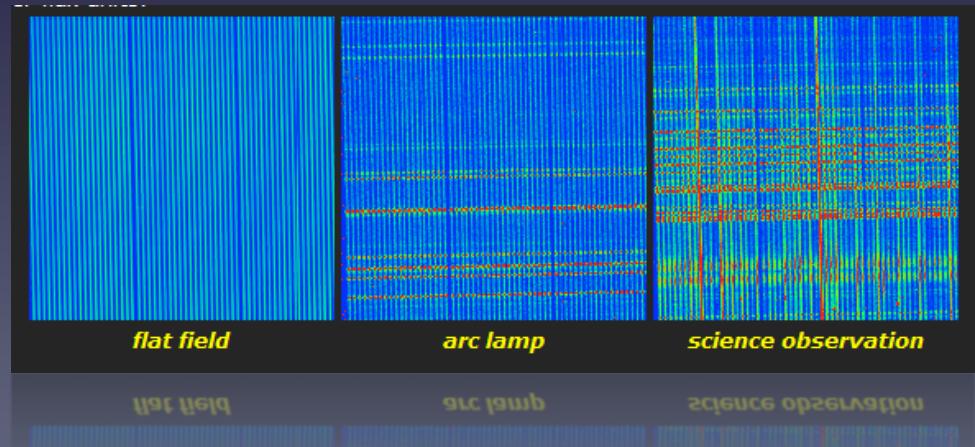
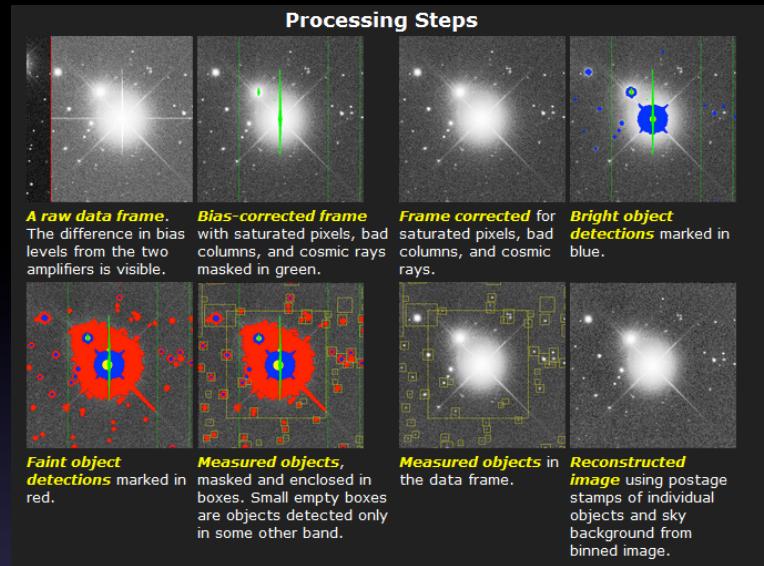
Spectroscopic Footprint

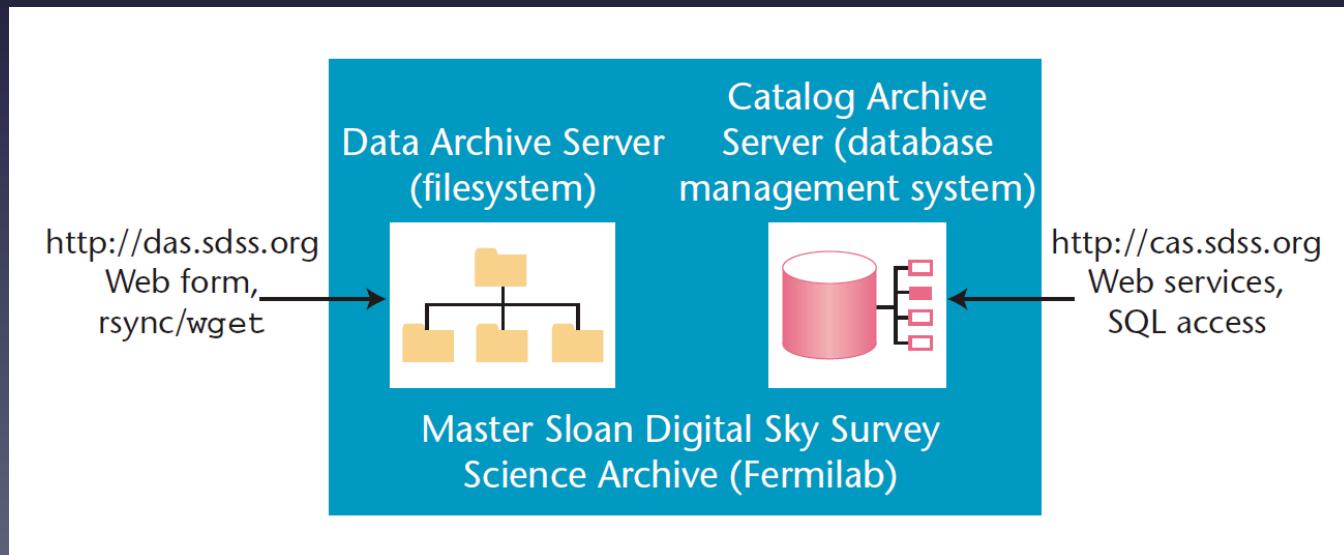
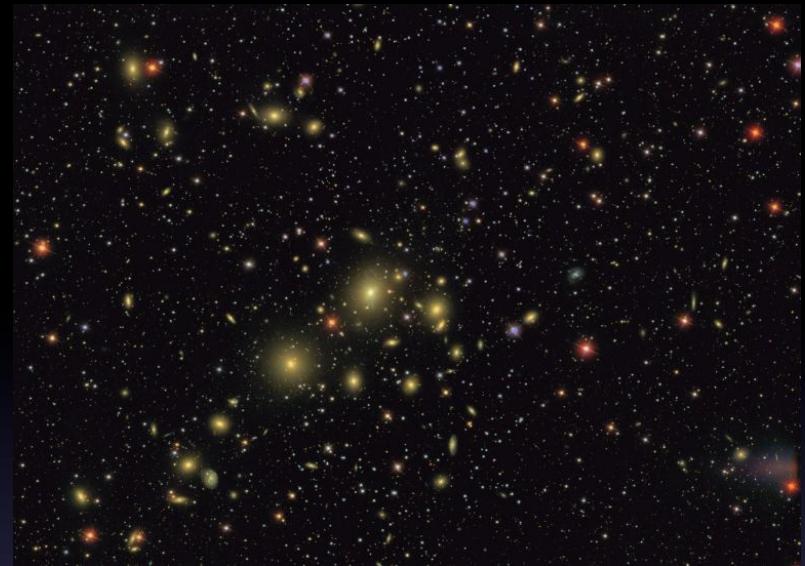
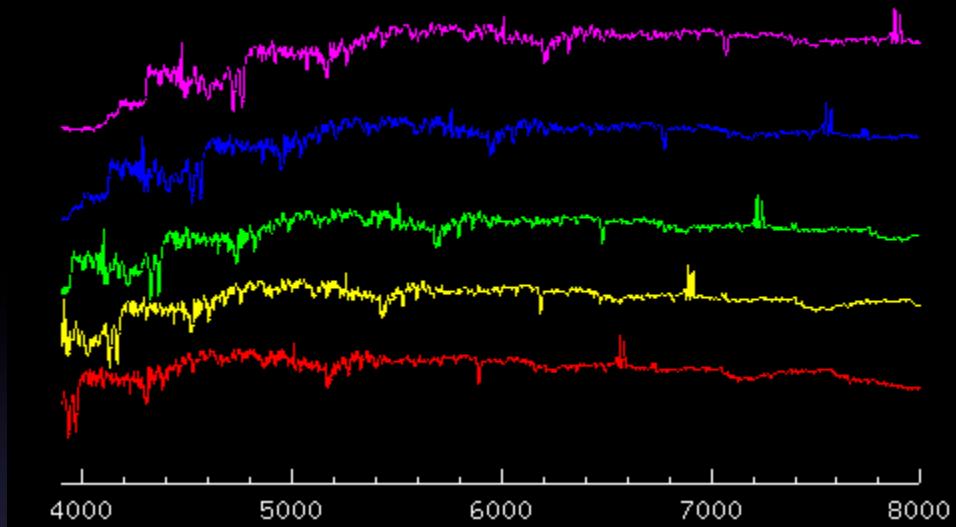


Spectrographs

- The original spectrograph
 - SDSS spectrograph (x2, optical, 640 fibers total, SDSS I & II)
 - Built for the original survey; a redshift-measuring machine
 - <http://skyserver.sdss.org/dr4/en/sdss/instruments/instruments.asp#spectrographs>
- SDSS III instruments
 - BOSS spectrograph (optical, 1000 fibers total, R=1560 (red) – 2270 (blue))
 - Upgrade of the original SDSS spectrographs
 - http://www.sdss.org/instruments/boss_spectrograph/
 - APOGEE spectrograph (infrared, 300 fibers, SDSS III, R~22,500)
 - Built for Milky Way archeology
 - http://www.sdss.org/instruments/apogee_spectrograph/
 - MARVELS spectrograph (optical, 60 fibers, SDSS III)
 - Built to search planets around nearby stars with radial-velocity measurements
 - http://www.sdss.org/instruments/marvels_spectrograph/

Data Processing on SDSS

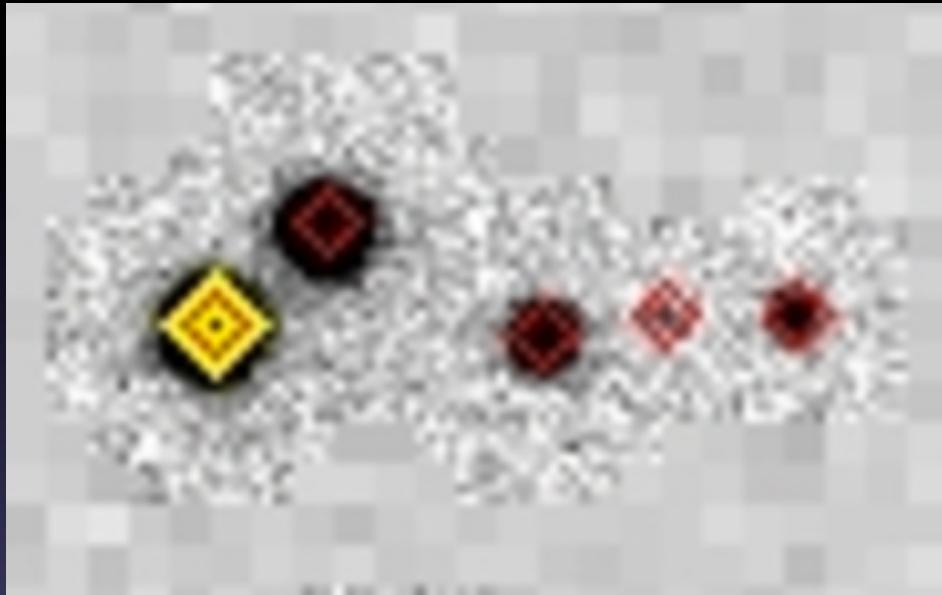




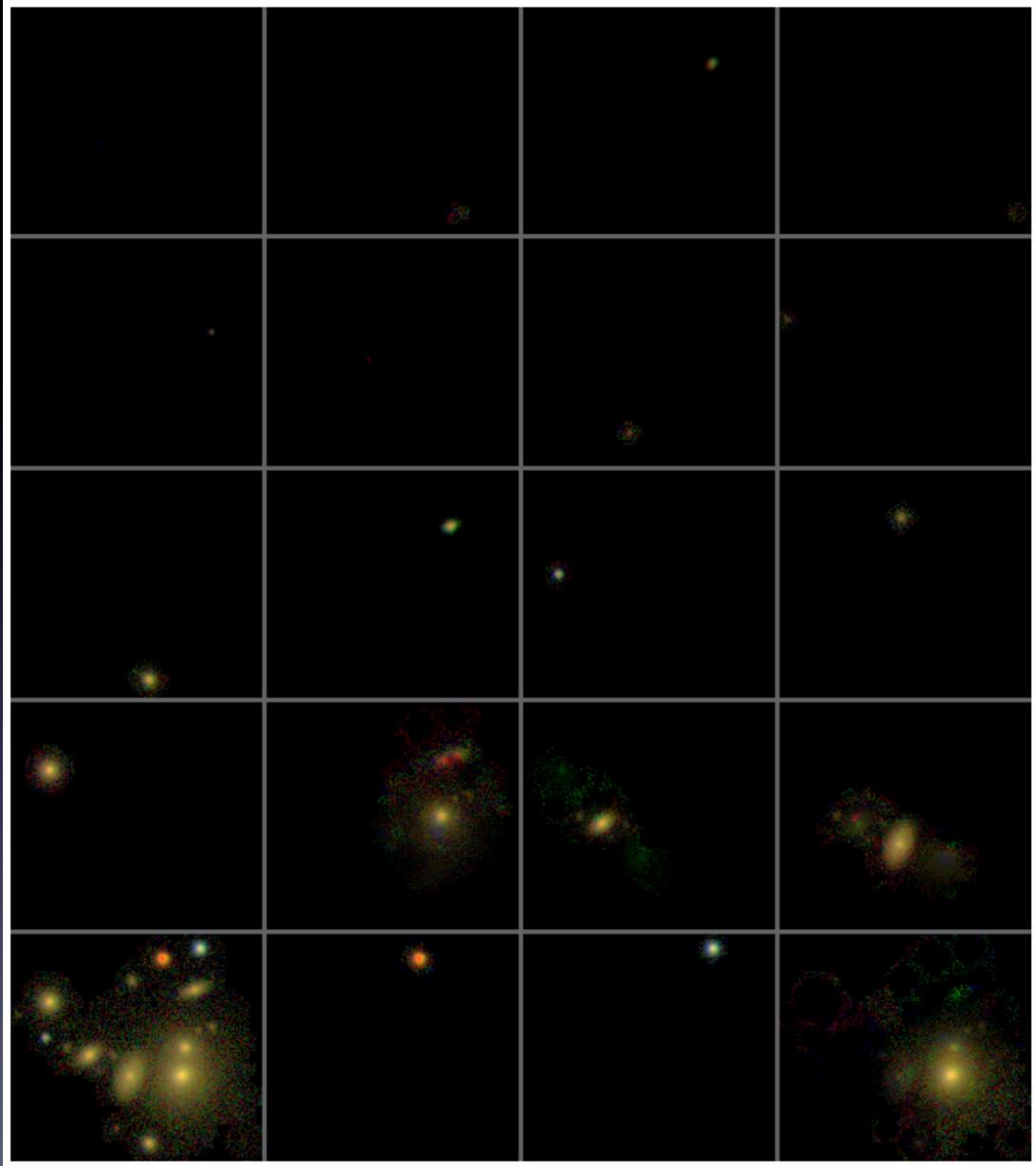
Imaging Pipelines

- All SDSS imaging data is processed with an automated software pipeline called *Photo*, written by Robert Lupton (Princeton) & Zeljko Ivezic (UW)
 - The still best reference for *Photo*: <http://www.astro.princeton.edu/~rhl/photo-lite.pdf>
- This pipeline (among other things):
 - Reduces the SDSS images (accounts for instrumental effects – dark frames, flat fields, bias, CCD defects – as well as removes nuisances such as cosmic rays)
 - Detects objects in each observed frame
 - Attempts to deblend blended objects
 - Measures (potentially deblended) object magnitudes (in a few different ways), and their shapes
- **This was the first pipeline that could do at scale this with minimal human supervision. Necessary at data rates that SDSS has achieved.**
- Other important imaging-related pipelines:
 - The *Resolve* pipeline: finds repeated observations of the same object
 - The *Ubercalibration* pipeline: global photometric calibration (i.e., ensure that 15th mag means the same thing no matter where we are on the sky)

Deblending



Left: Illustration of "parent" and "child" objects. This is a small patch of an image of the sky, with five individual astronomical sources. The photometric pipeline detects all five together as a single "parent" object, and determines the center of the parent to be at the yellow diamond. The deblending procedure then breaks up the parent into five children, whose centers are shown as the red diamonds.



Left: Unfortunately, the blends are often much more complex.

The bottom left panel shows the observed scene (all detected as a single object!), while the subsequent panels show the result of deblending. In this particular example, the deblender performed well; this is not always the case.

Deblending is a challenging problem.

(image courtesy of Robert Lupton)

Nanomaggies and *asinh* magnitudes

- SDSS stores fluxes in “*nanomaggies*”. An object with a flux of 1 maggy corresponds to 0th magnitude. Therefore, an object with flux of 1 nanomaggy (== 10⁻⁹ maggies) has a Pogson flux of -22.5 (which is close to SDSS flux limit). Remember $m = -2.5 \log(f/f_0)$.
- Things are a bit more complicated than that, though: Magnitudes within the SDSS are expressed as inverse hyperbolic sine (or “*asinh*”) magnitudes (see Lupton, Gunn and Szalay 1999):

$$m = -2.5/\ln(10) * [\text{asinh}((f/f_0)/(2b)) + \ln(b)].$$

- This was done to be able to deal with negative fluxes. In practical terms asinh magnitudes (also affectionately known as Luptitudes) are identical to Pogson magnitudes unless operating near the flux limit (~22 mag).

Spectroscopic Pipelines

- The original spectro pipeline
 - Measure the 1d spectrum and estimate redshift, spectral classification, and individual line strengths
 - Used in SDSS I/II/III (BOSS)
 - Described in Stoughton et al. (2002), but evolved since (and significantly updated for BOSS)
 - <http://www.sdss.org/dr12/spectro/pipeline/>
- The Stellar Parameters Pipeline (SSPP)
 - Post-processes the 1d spectra extracted with the SDSS spectroscopic pipeline to estimate stellar parameters (e.g., radial velocity, temperature, metallicity, gravity, alpha abundances)
 - Lee et al (2008abc) series of papers
- APOGEE pipelines (apred & aspcap)
 - Stellar parameters from APOGEE spectrograph and individual abundances of 15 elements
 - <http://www.sdss.org/dr12/irspec/apred/>
- MARVELS pipeline
 - Determining radial velocities
 - <http://www.sdss.org/dr12/marvels/pipeline/>

<http://www.sdss3.org/dr9/spectro/pipeline.php>

SDSS Data

- These pipelines produce a series of data products:
 - Reduced images, stored and available as FITS files
 - 1D reduced spectra, stored as FITS files
 - Catalogs, stored in SQL databases and FITS files
- Catalogs are typically the most “science-ready” product, but others are useful as well

Data Releases

- The SDSS Collaboration makes the data public in Data Releases, every 1-2 years
- Typically, in each data release all the data from previous data releases is reprocessed as well, to benefit from improvements in the software and the understanding of the telescope and instruments.
- There have been 13 data releases (including the Early Data Release in 2002). The current data release is DR 12 (more later)

The SDSS Surveys

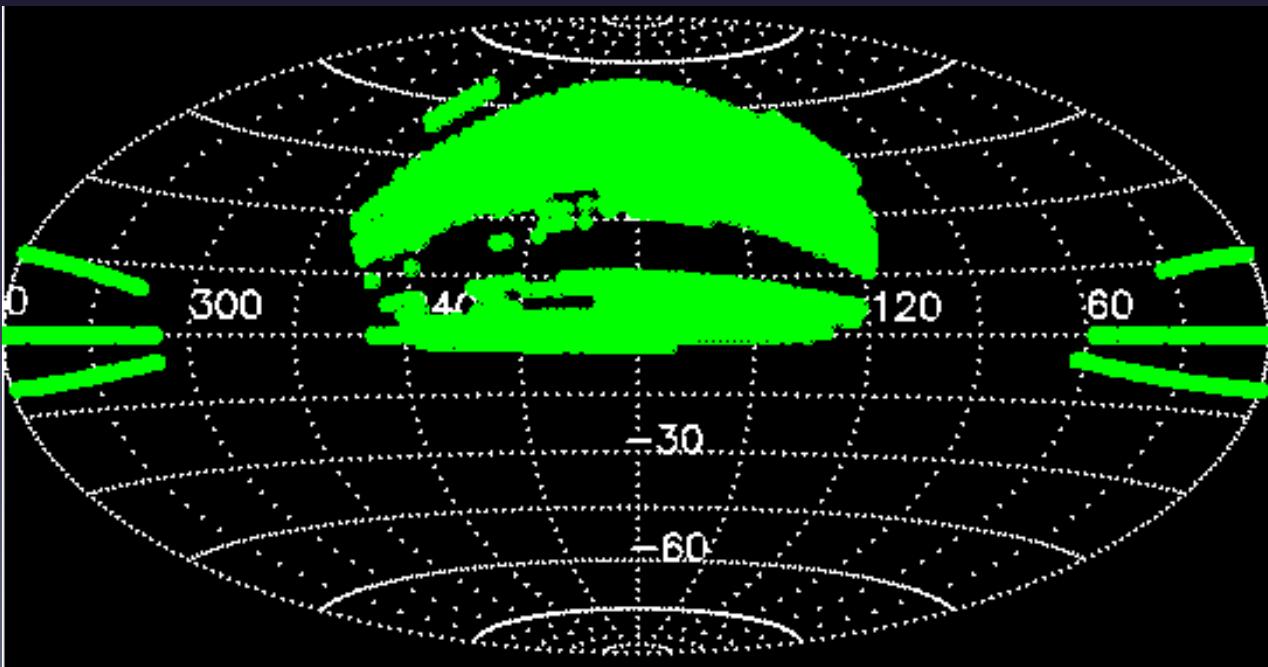
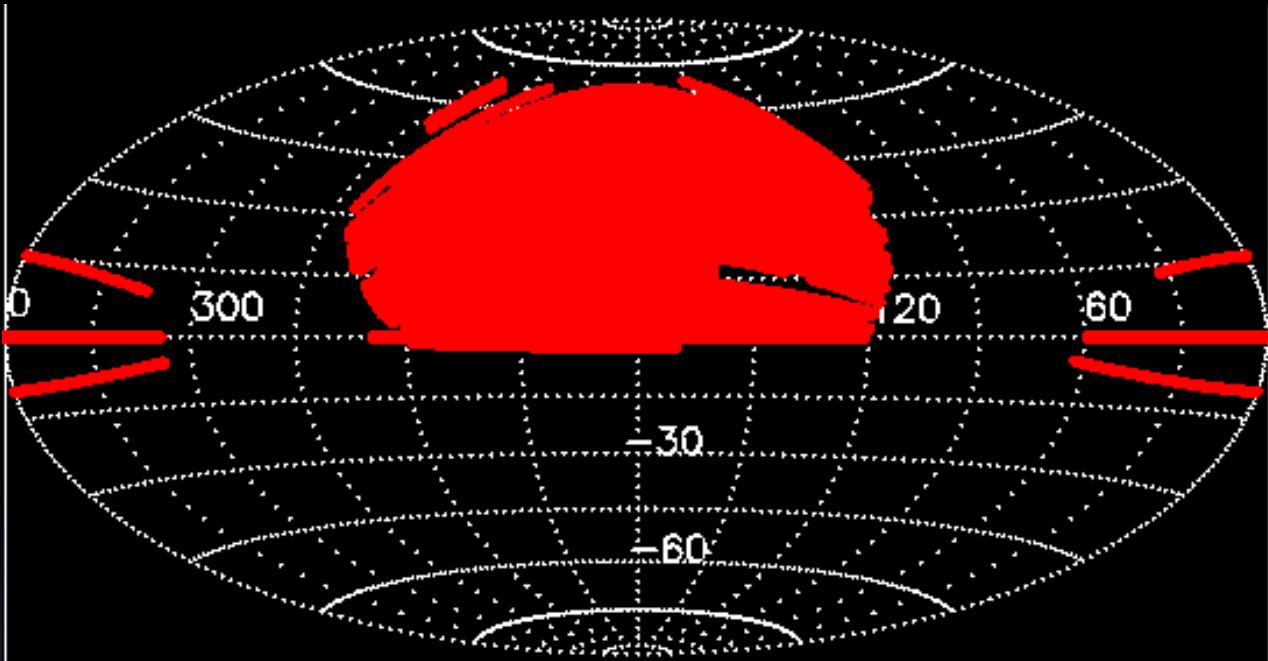
SDSS I (2000-2005)

- Built to conduct the largest redshift survey to date, with one of the primary goals being the measurement of the 3D galaxy power spectrum and constraining Λ CDM parameters.
- Imaging and spectroscopy
- Final data release: DR5
- Data access: <http://classic.sdss.org/dr5/>

Footprint area	8000 sq. deg.				
Imaging catalog	215 million unique objects				
Data volume	images			9.0 TB	
	catalogs (DAS , fits format)			1.8 TB	
	catalogs (CAS , SQL database)			3.6 TB	
Photometric calibration	<i>r</i> 2%	<i>u-g</i> 3%	<i>g-r</i> 2%	<i>r-i</i> 2%	<i>i-z</i> 3%

Spectroscopic area	5740 sq. deg.			
Wavelength coverage	3800-9200Å			
Resolution	1800			
Signal-to-noise	>4 per pixel at <i>g</i> =20.2			
Redshift accuracy	30 km/sec rms for main galaxy sample (from repeat observations)			
Target magnitude limits for main samples	Galaxies: Petrosian <i>r</i> < 17.77 Quasars: PSF <i>i</i> < 19.1			
Spectroscopic catalog	1,048,960 spectra, classified into 674,749 Galaxies 79,394 Quasars (redshift < 2.3) 11,217 Quasars (redshift > 2.3) 154,925 Stars 60,808 M stars and later 55,555 Sky spectra 12,312 Unknown class			

SDSS I (2000-2005)



Top: photometric
coverage

Bottom: spectroscopic
coverage

SDSS II (2005-2008)

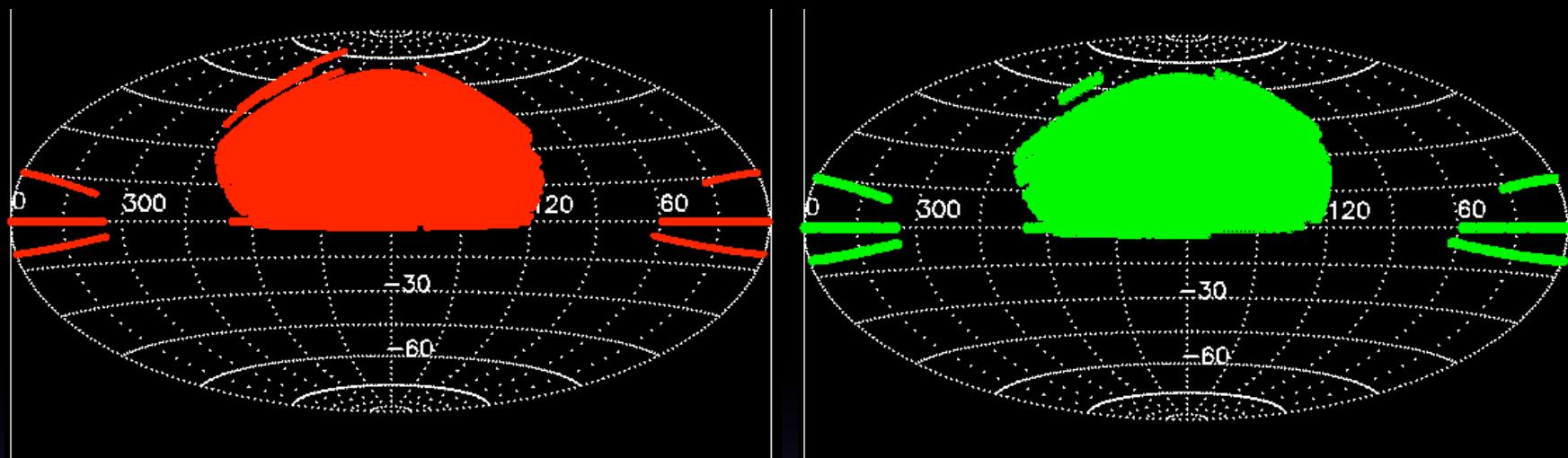
- Extension of SDSS II, with the goal to complete the SDSS I (“Legacy”) Survey, expand into Galactic structure research, and look for cosmologically interesting supernovae
 - Beginning of SDSS “subsurveys” (Legacy, SEGUE, Stripe 82)
- Imaging and spectroscopy
- Final data release: DR7
- Data access:
<http://classic.sdss.org/dr7/>

Footprint area	Total cataloged in CAS	11663 sq. deg.
	Legacy unique	8423 sq. deg.
	Legacy NGC ellipse	7646 sq. deg.
	SEGUE	3240 sq. deg.
	Supernova Survey	~300 sq. deg., repeated ~80 times
	M31 / Perseus / Sgr / SGP scans	46 sq. deg.
	Low galactic latitude fields ("Orion" runs)	832 sq. deg.
Imaging total area in DAS (multiple scans counted multiple times)	45,000 sq. deg. (1.3 million frames/filter)	
Imaging catalog	357 million unique objects (SEGUE: 127 million, Legacy: 230 million)	
Data volume	images (fits)	15.7 TB
	other data products (catalogs, masks, jpeg images, etc.) (DAS , fits format)	26.8 TB
	catalogs (CAS , SQL database)	18 TB

Spectroscopic area	Total	9380 sq. deg.
	Legacy	8032 sq. deg.
	SEGUE	1348 sq. deg.

Class	N(total)	N(main)	N(SEGUE)
All	1,640,960	1,374,080	266,880
Galaxies	929,555	928,567	988
Quasars ($z < 2.3$)	104,740	103,121	1,619
Quasars ($z \geq 2.3$)	16,633	15,411	1,222
M stars and later	84,047	76,125	7,922
Other stars	380,214	150,748	229,466
Sky spectra	97,398	75,209	22,189
Unknown	28,383	24,767	3,616

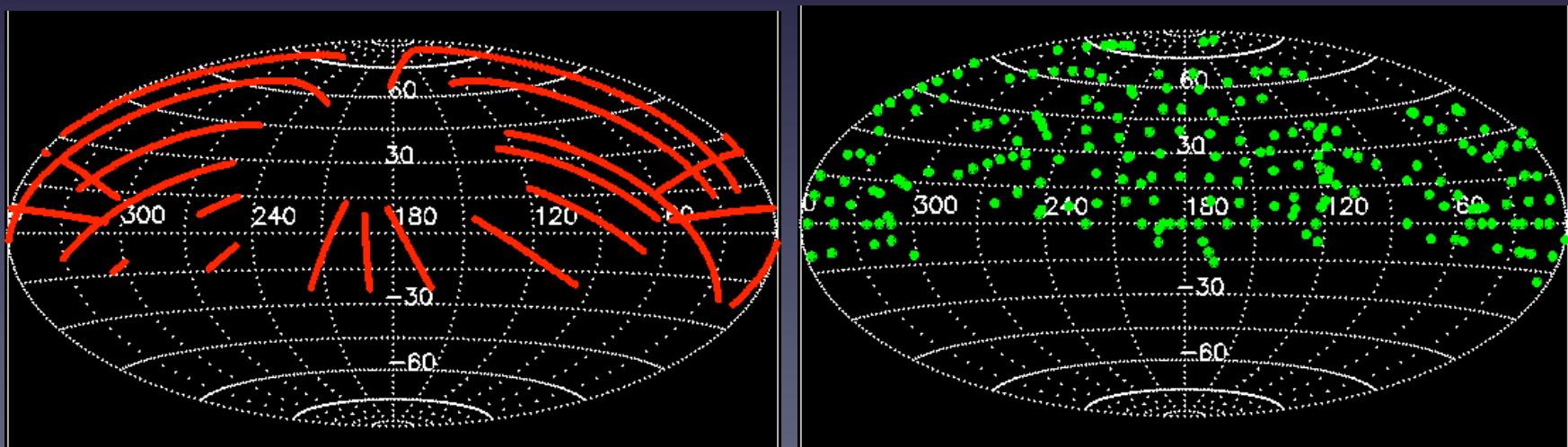
Spectroscopic catalog



Above: Photometric (right) and spectroscopic (left) coverage of SDSS II DR7 Legacy Survey

Not shown: ~80 epochs of repeated Stripe 82 imaging (Supernova Survey)

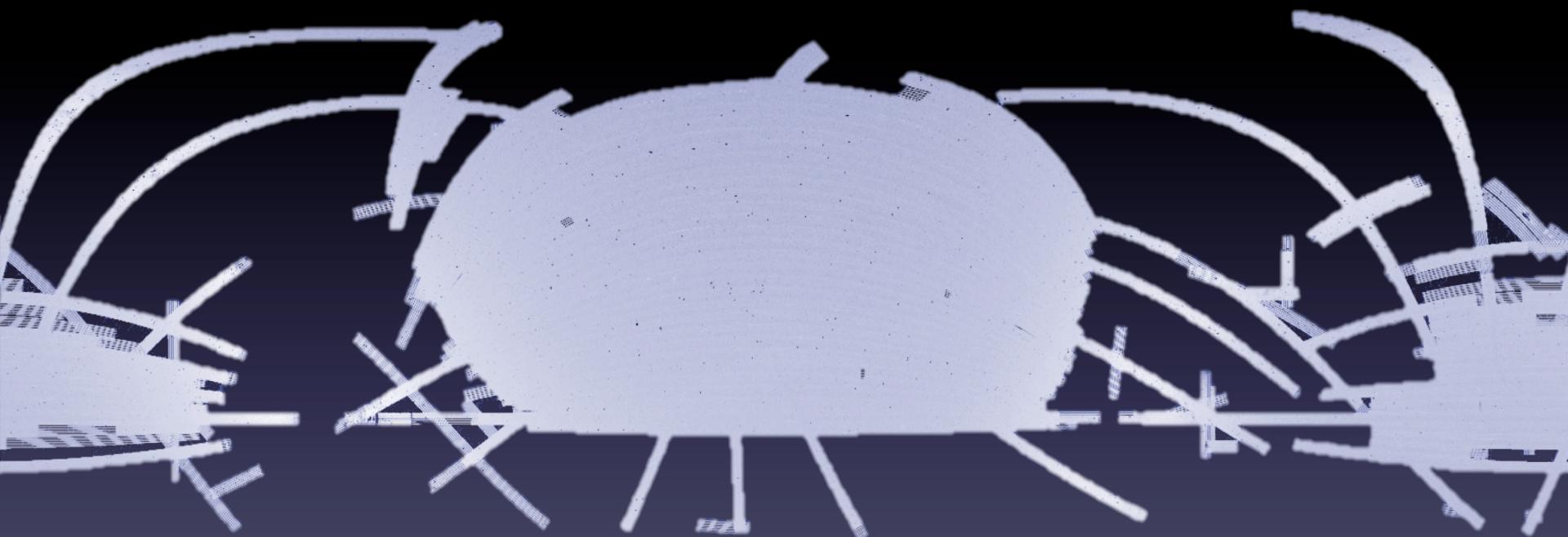
Below: Photometric (right) and spectroscopic (left) coverage of SDSS II DR7 SEGUE Survey



SDSS III (2008-2014)

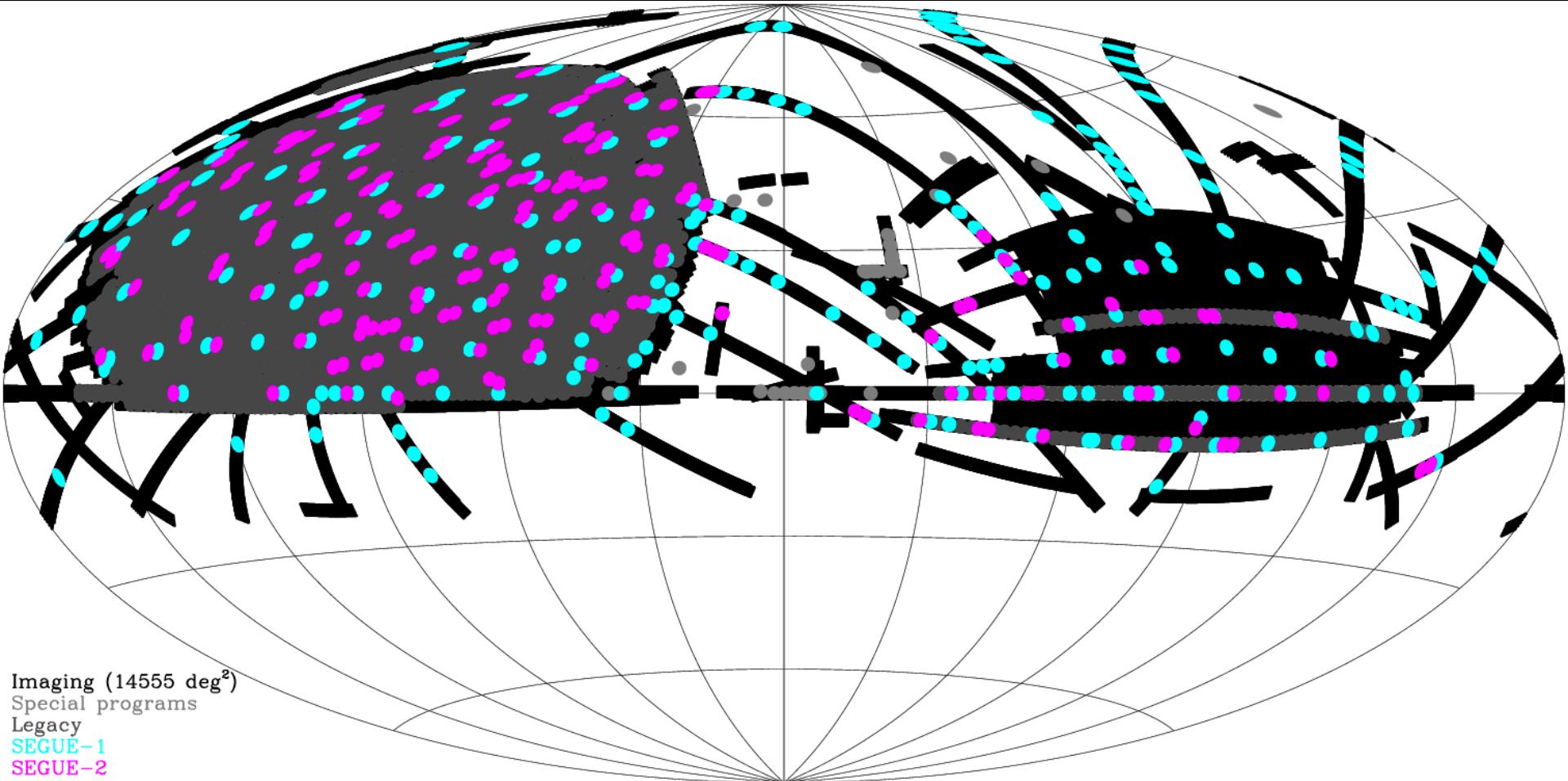
- Four distinct surveys conducted using the same telescope
 - APOGEE (the APO Galactic Evolution Experiment)
 - BOSS (Baryon Oscillation Spectroscopic Survey)
 - MARVELS (Multi-object APO Radial Velocity Exoplanet Large-area Survey)
 - SEGUE-2 (continuation of SEGUE)
- Largely spectroscopy only (the camera is now in the Smithsonian – last image was taken on November 18, 2009!)
- Final data release: DR12
 - Includes the SDSS I/II imaging as well
- Data access: <http://www.sdss.org/dr12/>

Final Photometric Dataset



SDSS Data Release 12 imaging footprint, encompassing 14500 deg².
Imaging was completed in DR8; no new imaging is being done.

DR12 Imaging Footprint and SEGUE I/II Spectroscopic Fields



Photometric Dataset Statistics

Total unique area covered 14,555 square degrees

Total area of imaging 31,637 square degrees
(including overlaps) (excluding supernova runs)

Individual image field size 1361×2048 pixels (0.0337 square degrees)

Number of individual fields 938,046 (excluding supernova runs)

Number of catalog objects 1,231,051,050

Number of unique detections 932,891,133

Median PSF FWHM, r -band 1.3 arcsec

Pixel scale 0.396 arcsec

Exposure time per band 53.9 sec

Time difference between observations of each band 71.72 sec (in *riuzg* order)

Global astrometric precision 0.1 arcsec rms (absolute)

Number of unique, primary sources

Total 469,053,874

Stars 260,562,744

Galaxies 208,478,448

Unknown 12,682

Effective wavelengths & magnitude limits
(95% completeness for point sources)

u 3551Å
22.0

g 4686Å
22.2

r 6165Å
22.2

i 7481Å
21.3

z 8931Å
20.5

Relative photometric calibration accuracy (RMS)
[\(Padmanabhan et al. 2008\)](#)

u 1.3%

g 0.8%

r 0.8%

i 0.7%

z 0.8%

APOGEE Targets

DR 12: APOGEE

- APOGEE survey used a high-resolution IR spectrograph to measure spectra of over 100,000 red giants and other stars. It measured:
 - their radial velocities
 - stellar parameters
 - and detailed abundances of 15 chemical elements to 0.1 dex
- Why: Galactic archeology
- <http://www.sdss.org/surveys/apogee/>

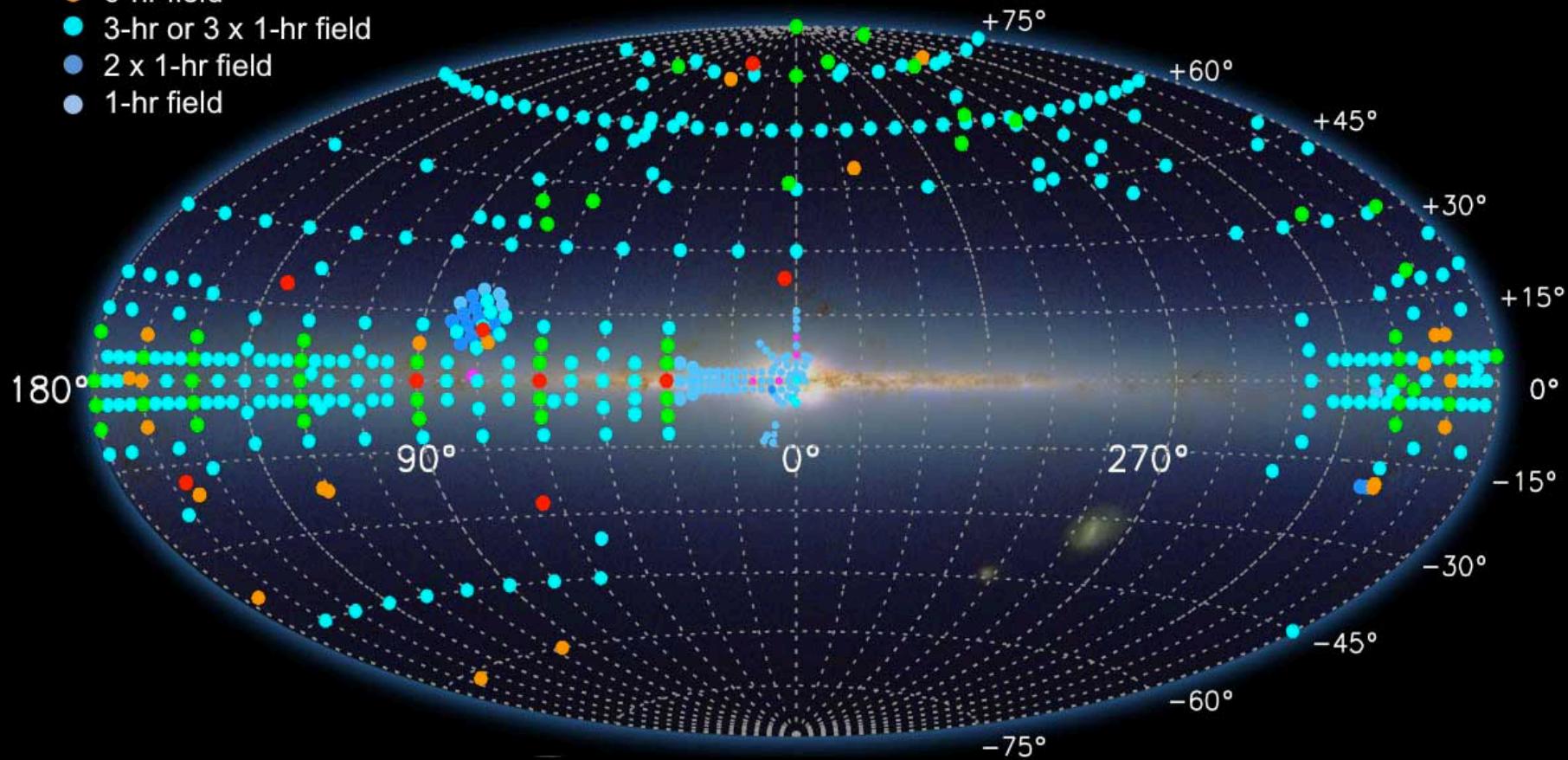
DR12 includes data for ~163,000 APOGEE targets. This includes 146,000 science targets, located in distinct types of survey fields:

~15,000 stars in Bulge fields
~28,000 stars in Halo fields
~55,000 stars in Disk fields
~14,000 stars in Kepler/CoRoT fields
~8,000 objects in Ancillary Science fields
~1,800 stars in Halo Stream fields
~1,200 stars in Sagittarius dSph fields
~8,000 stars in Star Cluster fields
~900 bright stars observed with the NMSU 1m telescope + APOGEE, including bright standards

The APOGEE main survey sample also includes ~17,000 hot stars used for telluric correction, across all field types.

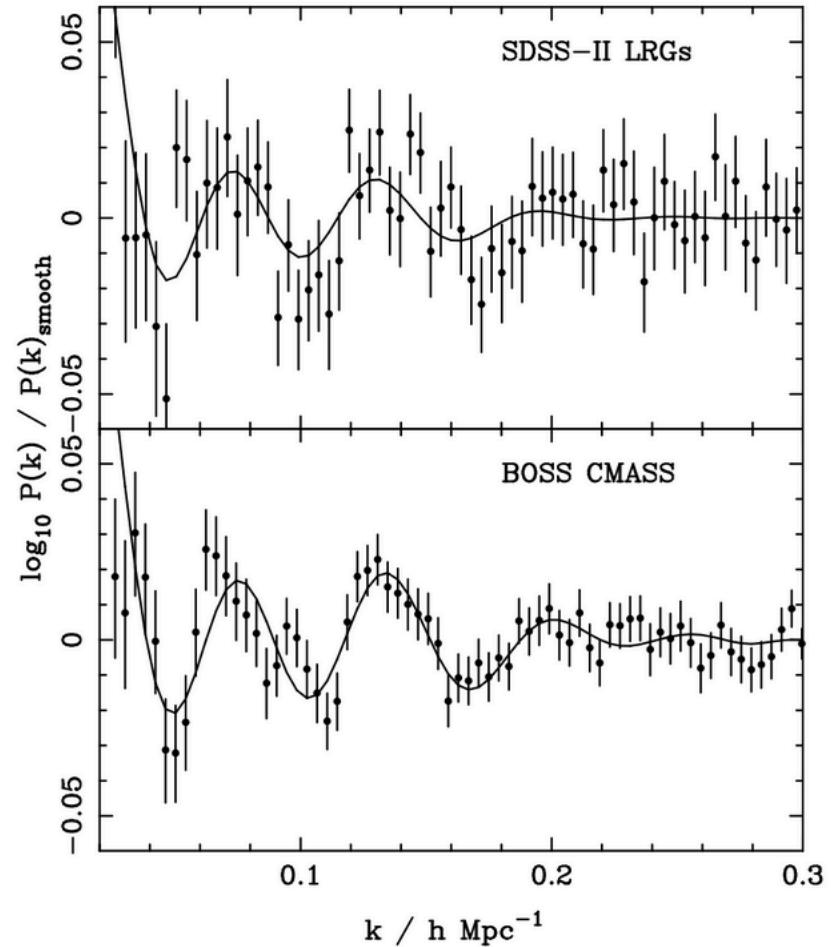
APOGEE DR12 Coverage - Observed Survey Plan

- Commissioning field – 1-hr
- 24-hr field
- 12-hr field
- 6-hr field
- 3-hr or 3 x 1-hr field
- 2 x 1-hr field
- 1-hr field



DR 12: BOSS

- BOSS: Baryon Oscillation Spectroscopic Survey
 - Designed to map the spatial distribution of luminous red galaxies (LRGs) and quasars to detect the characteristic scale imprinted by baryon acoustic oscillations in the early universe.
 - Do this by collecting spectra and measuring redshifts of 1.5 million galaxies out to redshift of $z=0.7$, and spectra of 160k quasars in $2.2 < z < 3$ range (for Lyman α forest)
 - <http://www.sdss.org/surveys/boss/>
- Upgraded original SDSS spectrographs to increase throughput and precision



Comparison of the power spectrum of SDSS-II LRGs and BOSS DR9 CMASS galaxies. Solid lines show the best-fit models. From Anderson et al. 2012.

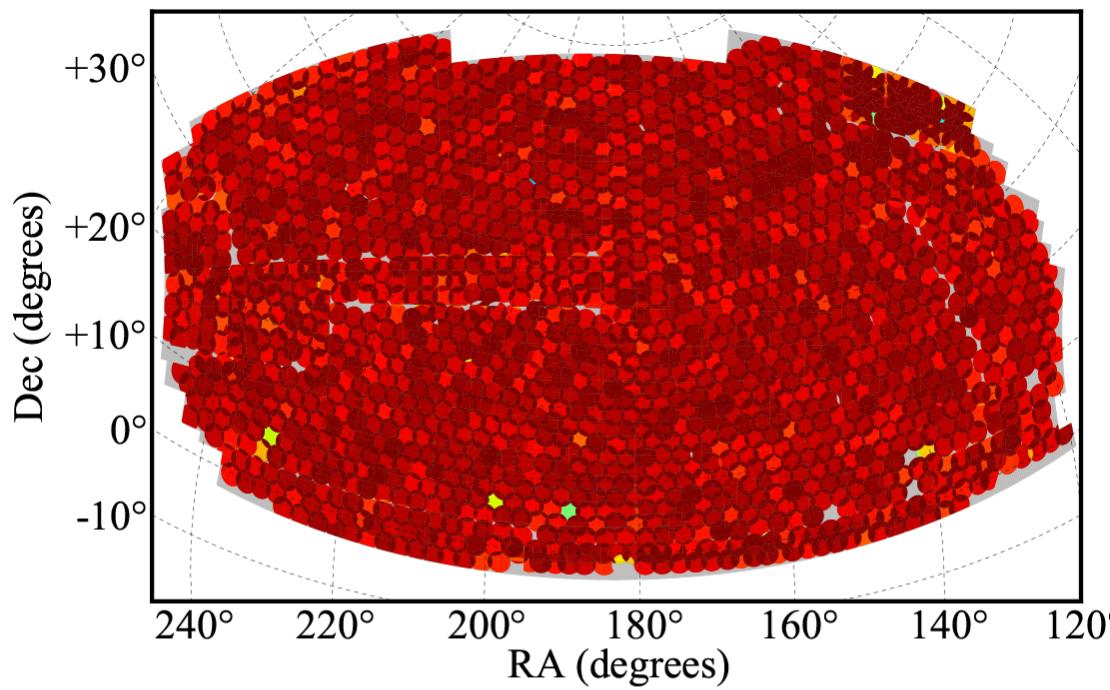
DR 12: BOSS

- Both the spectra and estimated parameters for BOSS observations are in SDSS DR12
 - “basic” parameters such as redshifts
 - Added value catalog of galaxy properties (e.g. stellar masses)
 - SSPP parameters for observed stars
- Note there are many stars (hundreds of thousands!) in the BOSS dataset

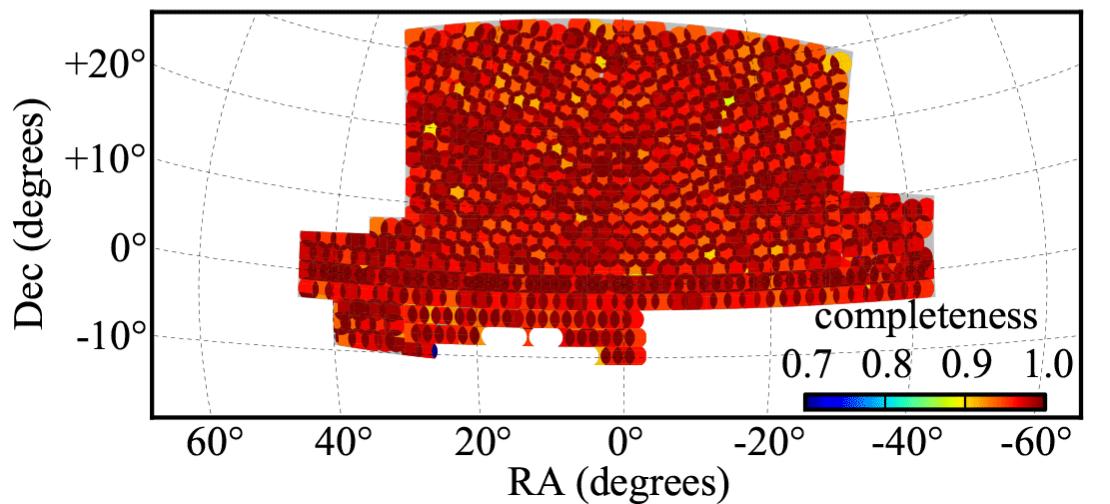
Extragalactic (BOSS)

Statistic	Total	Unique
Spectroscopic effective area (deg ²)	...	9,376
Plates	2,512	2,438
Spectra	2,497,484	2,269,478
All Galaxies	1,480,945	1,372,737
CMASS	931,517	862,735
LOWZ	368,335	343,160
All Quasars	350,793	294,512
Main	241,516	220,377
Main, $2.15 \leq z \leq 3.5$	175,244	158,917
Stars	274,811	247,216
Standard stars	52,328	42,815
Sky	238,094	223,541
Unclassified spectra	163,377	140,533

DR12

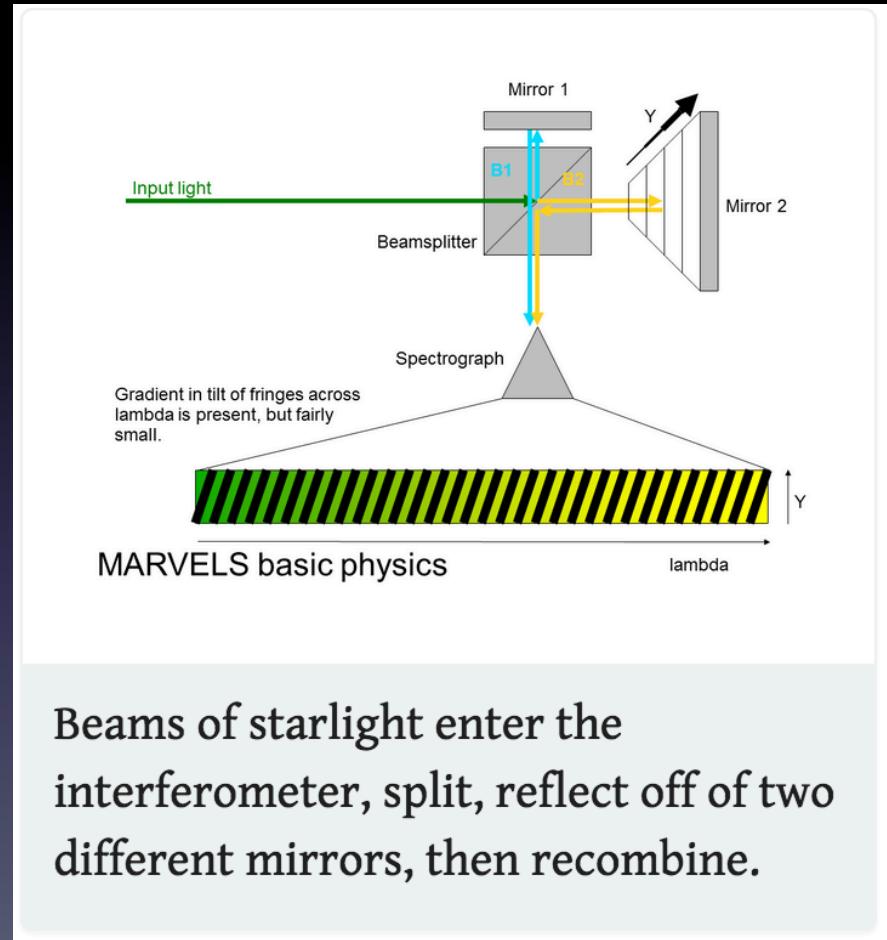


DR 12 BOSS coverage

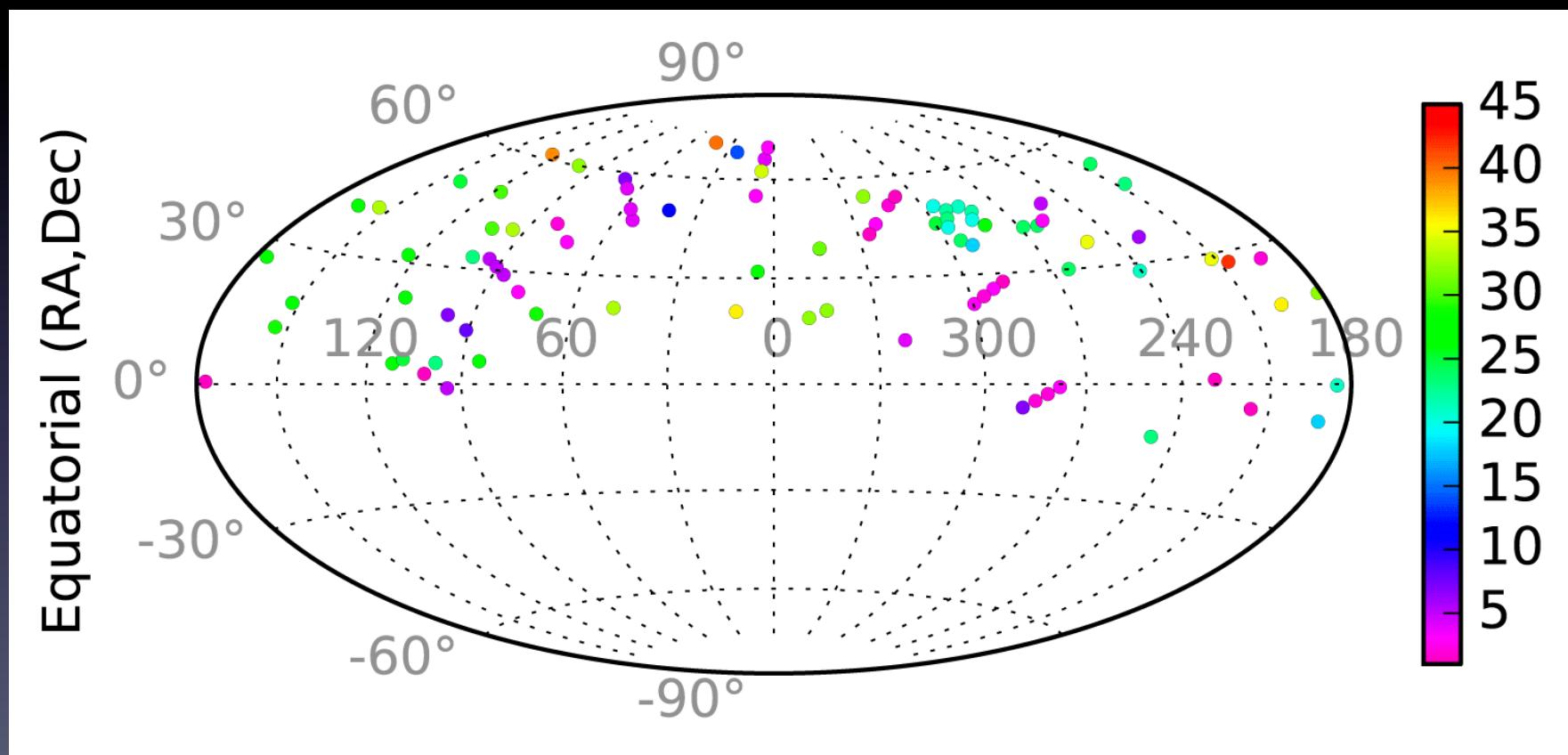


DR 12: MARVELS

- MARVELS: Multi-object APO Radial Velocity Exoplanet Large-area Survey
 - Observing from fall 2008-summer 2012, operating in bright time
 - Observed ~3500 stars ($7.6 < V < 12$)
 - 20-40 pointings per star
- Available data:
 - Radial velocity time series for each star
 - Reduced 1d spectra
- Note: this is a relatively small data set



DR 12 MARVELS Survey Coverage



Accessing the Data

Each Data Release includes four types of data: **images**, **optical spectra**, **infrared spectra**, and **catalog data** (parameters measured from images and spectra, such as magnitudes and redshifts).

The SDSS offers several different online data access tools, each suited to a particular need:

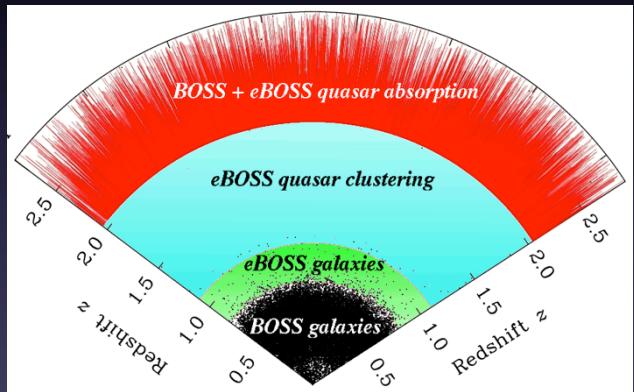
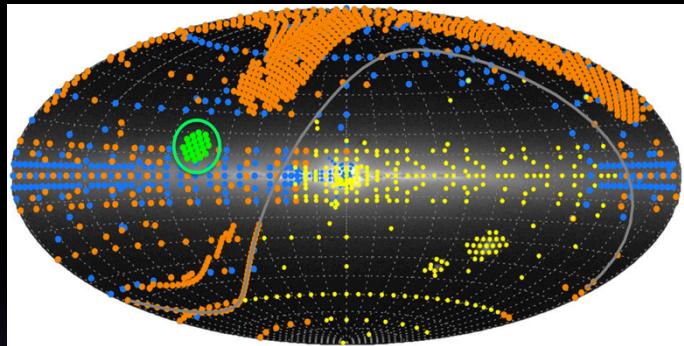
Website	Purpose
Science Archive Server	Interactive spectra and image mosaics
SkyServer	Browser-based access to the Catalog Archive Server (CAS) database, with resources for learning SQL and projects to teach science
CasJobs	Flexible advanced SQL-based interface to the CAS, for all data releases (quick registration required)
DR12 FITS	Direct download access to DR12 FITS data files for experts
Data Model	Details of the SAS directory structure, file formats, and the contents of each file

The data are available through the SQL database interface (see Lectures #1 and #4), various specialized interactive interfaces (e.g., spectrum viewers, image viewers), and downloadable in bulk (HW #1 Problem #3).

See <http://www.sdss.org/dr12/> for details.

The Future: SDSS IV (2014-2020)

- Three Surveys, 2014-2020
- APOGEE-2
 - Building and deploying another APOGEE-like IR spectrograph in the south (2.5m du Pont Telescope at Las Campanas Observatory in Chile)
- eBOSS (extended BOSS)
 - Continuation of BOSS
 - Time-Domain Spectroscopic Survey (TDSS)
 - Follow up time-variable targets
 - Note: PI is Scott Anderson
 - SPIDERS
 - Follow-up of eRosita X-ray sources
- MaNGA (Mapping Nearby Galaxies at APO)
 - An Integral Field Unit (IFU) survey to map the structure of nearby galaxies
 - 17 IFUs per pointing (19-127 fibers per pointing)
 - 10,000 galaxies
 - Note: Nick MacDonald (UW) is the chief engineer & project manager; the instrument is largely being built here



The Legacy of SDSS

- Huge science legacy:
 - The discovery of the most distant quasars (Fan et al. 1999)
 - First detection of the Baryon Acoustic Oscillation peak in distribution of galaxies (Eisenstein et al. 2005)
 - Measurement of Galactic structure parameters to unprecedented precision (Juric, Ivezić, Bond, etc...)
 - Discovery of ubiquitous Milky Way substructure (Newberg, Belokurov, Juric, Ivezić, etc..)
 - Discovery of hypervelocity stars (Brown et al)
 - The largest catalog of asteroid colors, evidence for genetic relationship in asteroid families, evidence for space weathering, ... (Ivezić, Juric, Jedicke, etc..)
 - ...
- More than 2000 refereed articles, with more than 70,000 citations!
 - This is a statistic from 2008! (<http://classic.sdss.org/signature.html>)

The Legacy of SDSS

- Equally important is its impact on how we do astronomy:
 - Data Driven Astronomy: The project that unambiguously showed the broad capabilities of large sky surveys, for uses beyond their original intent
 - Software Instrumentation: The first survey where software was equally important (and difficult) as the hardware and the instruments
 - Large Collaborations: The project that brought the age of large collaborations into astronomy
 - Open Data: The survey that showed the value of public (and frequently released) data

Summary: The SDSS Surveys

- SDSS: 2000-2005
 - Imaging and spectroscopic survey (8000 deg², 5700 deg²)
- SDSS II: 2005-2008
 - Imaging and spectroscopy
 - Legacy, SEGUE, and Supernova Search programs
- SDSS III: 2008-2014
 - APOGEE (the APO Galactic Evolution Experiment)
- BOSS (Baryon Oscillation Spectroscopic Survey)
- MARVELS (Multi-object APO Radial Velocity Exoplanet Large-area Survey)
- SEGUE-2
- SDSS IV: 2014-2020
 - APOGEE-2
 - eBOSS (extended BOSS)
 - MaNGA (Mapping Nearby Galaxies at APO)