Sloan Digital Sky Survey
Five-Year Baseline Plan for SDSS Operations

Revised February 26, 2001

Introduction

Our goal is to survey pi steradians of the North Galactic Cap and obtain one million galaxy redshifts and 100,000 quasar redshifts. All of the systems have been designed to meet this goal. To forecast the rate of covering the sky, we have devised a model that is intended to capture the key limiting factors. We have carried the projection out to a five-year survey duration, which we call the Five-Year Baseline Plan, or “baseline” for short. The baseline has proven to be a useful tool for evaluating our progress to date. More than five years will be required to reach the ultimate goals and this will require additional funding beyond what the ARC Advisory Council has approved the project to seek.

The original baseline plan for the five-year survey was completed on April 1, 2000 and was included as the Appendix to the Quarterly Report: January 1, 2000 to April 11, 2000 (dated June 7, 2000). We stated in the baseline plan: “As we gain more experience in survey operations and can refine the elements of this projection (the baseline plan), we will revise the plan accordingly.” As of January 4, 2001, the Operations phase of the Survey has been underway for nine months and with that experience, we have revised the way the baseline plan is presented. There are two reasons for this. First, we need a more complex model to track the Southern Survey, which by itself is actually two distinct surveys, and second, we need to properly account for the overlaps when we convert from hours of observations to effective square degrees of imaging.

Northern Survey

The progress of the SDSS is being evaluated with respect to a projection that accounts for the available observing time and how efficiently we use it. The baseline for the five-year survey was derived from a model for the survey of the Northern Galactic Cap that takes as input the number of dark hours in each quarter, the relative allocations of that time to imaging and spectroscopy, and various efficiency factors. The model provides the total hours spent collecting data separately for imaging and spectroscopy. These two numbers can then be converted into square degrees and plates, which allows a straightforward comparison with the data actually collected.

It is important to know how the imaging survey is done in order to appreciate both the nomenclature and the complications. Area of sky is covered by scanning along a pre-defined set of great circles. A particular observation is called a “run”: it is a part of one of the great circles, the beginning and end are determined by the date of observation, the specific time during the night the atmosphere permitted good seeing, etc.

In more detail, the camera consists of 6 columns of detectors, each of which scans a narrow swath of sky called a scanline. Each scanline is 2048 pixels wide; since a pixel subtends 0.4 arcsec perpendicular to the scan direction, a scanline is 13.6 arc minutes (0.23 deg) wide. The scanlines are separated by gaps that are 90% of the width of a CCD detector, such that a second run can fill in the gaps with a bit of margin to spare. (This margin is needed because the telescope tracking is done open-loop.) A set of six scanlines is called a strip; the interlaced pair of strips is called a stripe. A stripe is a bit over 2.5 degrees wide. Having both strips of a stripe is necessary for the subsequent selection of spectroscopic targets, which means that the area covered by the stripes at any point in the survey is the “coin of the realm.”

Runs will be overlapped to avoid leaving holes, and the amount of repeat scanning needs to be accounted for. Eventually the whole of the great circle within the footprint is covered. Since runs will rarely be complete segments of a great circle.

Just as the scanlines overlap by about 10% of their width, so too the stripes overlap by a few arc minutes for the same reason: to allow for margin in the telescope pointing and tracking. The overlap between neighboring stripes varies depending on position along them because they are each centered on a great circle: the system of great circles of course converges at the poles. The minimum overlap is at the equator of the system of great circles; we defined the system such that stripes are separated by exactly 2.5 degrees there.

In summary, while the camera actually observes 20.5 square degrees per hour in terms of pixels on the sky, the effective rate of covering sky within a 2.5 degree-wide stripe is 9% less, or 18.75 square degrees per hour. The details of stripe-to-stripe overlap depend on the final geometry for the survey footprint, i.e. how close to the poles each stripe gets. The net overlap due to convergence of the stripes towards the poles is about 25%. The practical consequence is that to cover the whole area we need to scan about 25% longer than would be suggested by the formula 10,000 deg^2/18.75 deg^2/hr=533 hours. In the end, the imaging survey will cover a footprint of sky which is elliptical in outline and which avoids obscured directions in the Milky Way.

Not all of the imaging data obtained meet the quality standards of the survey - some data are rejected because the image quality is sub-standard, or because there is evidence that the sky was not uniformly transparent, or for some other reason. Thus as the data move through the pipelines, some data are rejected and that part of the sky becomes a candidate for re-observation.

To capture the nature of the way the data are collected and processed, we have devised a hierarchy of categories, each of which is tracked. These categories have the following definitions:

1. Gross – Total area scanned reported in idReport files. This is the sum of the area of the science runs that are good enough to be candidates for data processing. In the context of SDSS data processing, this should be the area that one computes by scanning all the idReport files in the "golden directory."
2. Net – "Total" area that has been successfully processed through the image pipelines. The excluded area includes ramp-up frames and the frames that were rejected due to bad tracking, clouds, setup time, or bad seeing. Net area is the area of sky stored in the Operations Database.
3. Good – "Net" area that passes QA tests for seeing, tracking, and photometricity. This statistic is not reduced by overlaps.
4. Unique – "Good" area that is corrected for overlaps between separate runs on the same stripe scans and that lies within the official survey area. This is one of the statistics that is used to track progress against the baseline and it was the statistic that was used to define the baseline of April 2000.
5. Footprint – "Unique" area after removing overlap of stripes due to the convergence of the great circles. This is ultimately the area of sky that the survey has covered and it is the statistic for the revised baseline that most realistically measures progress toward the survey goals for the North Galactic Cap and the southern outrigger stripes.
6. The repetition factor – The ratio of the sum of the Good areas scanned in a given footprint to the area of the footprint. It is one of the two statistics that are needed to measure progress toward the survey goals for the Southern Equatorial stripe. The other is the footprint statistic.

The baseline projection for sky area covered as a function of time refers to the category "Unique"– this is natural since the specifics of the stripe-to-stripe overlap is unknown to the baseline calculation. Thus the category "Unique" is useful for evaluating our operational efficiency, whereas the category "Footprint" is ultimately what counts.

Northern Survey Baseline

The model is presented in Table 1 and the number of dark hours listed in Table 1 is an approximate average for each quarter of each year for which the northern survey region is available. The values depend on decisions about 1) minimum horizon distance of Sun (nominal -18 deg); 2) minimum horizon distance of Moon (nominal 0 deg); 3) maximum airmass (nominal 2.2); and 4) minimum interval of time for observing on a given night (nominal 3 hrs). The hours for Q3 are low because of the frequency of thunderstorms in July and August and the fact that the Northern Galactic Cap is not visible from APO in September. The observatory will undertake routine maintenance during Q3, and typically, observations will not be scheduled during one of the dark periods during July and August. (In the model, we assume that it will be split between the July and August dark runs). In addition, we will generally schedule the re-aluminization of the primary mirror of the 2.5-m telescope during October since it cannot be done during July and August due to the demand for these facilities by Kitt Peak telescopes. We will inevitably lose part of a dark run between October and November and have adjusted the Q4 hours for this.

Table 1. SDSS 5-Year Model Projection – Northern Galactic Cap

	dark hrs.	im. frac.	sp. frac.	weath	uptime	im. eff.	sp. eff.	im. ops.	sp. ops.	end game	im. hrs.	sp. hrs.	sq. deg.	No. plts
2000 Q2	277	0.3	0	0.6	0.7	0.86	0.65	0.25		1	7.5	0	140.7	0
Q3	40	0.3	0	0.5	0.8	0.86	0.65	0.5		1	2.1	0	38.7	0
Q4	196	0.3	0	0.6	0.9	0.86	0.65	0.75		1	20.5	0	384.0	0
2001 Q1	464	0.3	0.7	0.6	0.9	0.86	0.65	0.9	0.75	1	58.2	85.5	1090.9	114
Q2	277	0.3	0.7	0.6	0.9	0.86	0.65	0.9	0.9	1	34.7	61.3	651.2	82
Q3	40	0.3	0.7	0.5	0.9	0.86	0.65	0.9	0.9	1	4.2	7.4	78.4	10
Q4	196	0.3	0.7	0.6	0.9	0.86	0.65	0.9	0.9	1	24.6	43.3	460.8	58
2002 Q1	464	0.3	0.7	0.6	0.9	0.86	0.65	0.9	0.9	1	58.2	102.6	1090.9	137
Q2	277	0.3	0.7	0.6	0.9	0.86	0.65	0.9	0.9	1	34.7	61.3	651.2	82
Q3	40	0.3	0.7	0.5	0.9	0.86	0.65	0.9	0.9	1	4.2	7.4	78.4	10
Q4	196	0.2	0.8	0.6	0.9	0.86	0.65	0.9	0.9	1	16.4	49.5	307.2	66
2003 Q1	464	0.2	0.8	0.6	0.9	0.86	0.65	0.9	0.9	1	38.8	117.3	727.3	156
Q2	277	0.2	0.8	0.6	0.9	0.86	0.65	0.9	0.9	1	23.2	70.0	434.2	93
Q3	40	0.2	0.8	0.5	0.9	0.86	0.65	0.9	0.9	1	2.8	8.4	52.2	11
Q4	196	0.2	0.8	0.6	0.9	0.86	0.65	0.9	0.9	1	16.4	49.5	307.2	66
2004 Q1	464	0.2	0.8	0.6	0.9	0.86	0.65	0.9	0.9	1	38.8	117.3	727.3	156
Q2	277	0.2	0.8	0.6	0.9	0.86	0.65	0.9	0.9	1	23.2	70.0	434.2	93
Q3	40	0.2	0.8	0.5	0.9	0.86	0.65	0.9	0.9	1	2.8	8.4	52.2	11
Q4	196	0	1	0.6	0.9	0.86	0.65	0.9	0.9	1	0.0	61.9	0.0	83
2005 Q1	464	0	1	0.6	0.9	0.86	0.65	0.9	0.9	1	0.0	146.6	0.0	195
Q2	277	0	1	0.6	0.9	0.86	0.65	0.9	0.9	1	0.0	87.5	0.0	117
TOTAL	5162										411.0	1155	7706.9	1540

In the baseline plan of April 2000, the spectroscopic survey was expected to start in Q1 of 2001, owing to the need for a head start for the imaging survey. Since the collection of imaging data, which met the survey requirements for target selection, actually began in the spring of 1999, it was possible to acquire spectroscopic data during the commissioning of the spectrographs. Some of this spectroscopic data has met survey requirements and it is included in the status reports. Even though this head start allowed spectroscopy to begin somewhat earlier than the baseline forecast, the baseline has not been changed. The 0.3 of the available dark time through Q2 of 2002 reflects the intent to front-load imaging (assuming, of course, that atmospheric conditions permit this relatively high fraction). The baseline plan assumes that imaging ends before spectroscopy is completed and thus spectroscopy has 100% of the available time beginning Q4 of 2004. In practice, the proportions will be adjusted as the survey progresses to yield the best strategy for maximizing both the imaging and the spectroscopy survey in five years.

The "weather" column shows the fraction of time that the weather conditions (clouds, seeing) can support at least spectroscopic observations. Based on historical records for 3.5-m telescope operations at APO, 60% of the dark time was astronomically useful in all months except for the monsoon months of July and August. In those two months, 40% of the time is astronomically useful. In the baseline, we assume that twenty percent of the astronomically useful time was photometric with good seeing. Good seeing is defined to be 1.5 arcseconds or better.

Uptime is the fraction of the observing time that the hardware and the mountaintop software are working well enough to enable data collection. The steady-state value of 90% is a reasonable goal given our experience so far.

The next two columns refer to the efficiencies of data collection, separately for imaging and spectroscopy. Inevitably some time will be lost due to the ramps for imaging, focusing, calibration exposures for spectroscopy, etc. These intervals of lost time can be estimated relatively precisely ahead of time, and these efficiencies are entered in the "imaging efficiency" and "spectroscopic efficiency" columns.

The imaging efficiency is estimated as follows. Early simulations suggested that a night of imaging would have an average length of 5.8 hours and would typically include two runs. Our experience is at least roughly consistent with this expectation. The ramps require about 8 minutes at the beginning and end of each run, and time must also be allocated to focus and set the rotator, say 15 minutes per run. Allowing the setup for the first run to have been concluded in the twilight, the best possible efficiency is then 0.86.

Spectroscopic efficiency assumes 25 minutes lost to calibrations, cartridge exchange, and field acquisition for each 45 min plate exposed. That's an efficiency of 0.65. In April 2000, we were meeting our goal of exchanging a cartridge in less than 10 minutes.

Other time is lost due to conditions that are in principle within our control, and these represent places where we can devise methods to enhance the use of telescope time. Estimates of these factors that affect how much time we expose on the sky are entered in the columns headed "imaging operations" and "spectroscopic operations." Examples of such factors include the loss of time at the end of a night because a run has finished and insufficient time remains to do anything else; exchanges between imaging and spectroscopy requiring additional time; and the time required to make a decision concerning a switch to spectroscopy if the seeing deteriorates during imaging. The values shown here, separately for imaging and spectroscopy, indicate that as the work progresses we will become more expert in minimizing such losses.

The column labeled "end game" anticipates that late in the survey, there will be times when none of the unmapped survey region is available, and we will cover no new area. Currently this is set equal to 1.0 in all quarters to reflect our intention to optimize the strategy for data collection.

These factors allow the hours spent in imaging and in spectroscopy to be computed. The CCD camera scans at an effective rate of 18.75 square degrees per hour in the course of producing a stripe. The number of square degrees observed is therefore obtained by multiplying the number of hours of imaging by 18.75. Similarly, the net spectroscopic time refers to the actual time integrating on the sky. The number of plates is obtained in the model by dividing the hours available for spectroscopy by 0.75 hours per plate.

After adopting these choices for the efficiencies and other entries, Table 1 shows that 7700 square degrees will be scanned in five years. (The full survey region of pi steradians, or 10,000 square degrees, actually requires 12,500 square degrees to be scanned, including overlapped area.) Simulations of the adaptive tiling suggest that 2100 plates are needed; the model predicts that 1540 plates can be exposed, which is 73% of the goal. The fraction between imaging and spectroscopy is pretty well balanced.

As we gain more experience in survey operations and can refine the elements of this projection, we will revise the plan accordingly. For the moment, the expected progress shown here serves as a point of departure for planning, and for consideration of the impact on the core scientific goals should the five-year run time be insufficient to cover the full survey area. Such a situation is characteristic of ground-based optical astronomy: the uncertainty caused by an incorrect projection of weather trends can be as significant as any failure to predict operational efficiencies.

Southern Survey Baseline

In the months of September, October, and most of November, the North Galactic Cap is not accessible and we concentrate instead on the South Galactic Cap - the Southern Survey. We show two tables for the Southern Survey: Table 2 describes the two "outrigger" stripes and Table 3 describes the equatorial stripe. This elaboration is motivated to recognize that these are, in fact, distinct surveys with separate strategies.

Outrigger Survey

Two stripes called "outriggers" flank the equatorial stripe. One is 108 degrees long (270 square degrees) and the other is 82 degrees long (205 square degrees). These outrigger stripes are to be observed in a way similar to the stripes in the Northern Survey, i.e., imaged once and targeted for spectroscopy with the same selection criteria. The total footprint is 475 square degrees. The combined length of the two outriggers is 190 degrees; with plates separated by typically 2.0 degrees along each of the two stripes, it will thus require 95 plates to finish the southern outrigger survey.

Table 2 for the outrigger survey was constructed by taking two-thirds of the available time for Q3, Q4 of 2000 and Q3 of 2001, then one-third of the time for Q4 2001, and zero time thereafter. The allocation between imaging and spectroscopy is the same as for the Northern Survey baseline after Q3 2002. All of the other factors are the same as for the Northern Survey baseline. This model yields a projection that by the end of 2001, we expect to have 650 square degrees of imaging and 112 plates, which are comfortably larger than 475 square degrees and 95 plates, respectively. Hence it is realistic to expect that we will have finished the southern outrigger survey by the end of 2001.

Table 2. SDSS 5-Year Model Projection – South Outrigger Survey

	dark hrs.	im. frac.	sp. frac.	weath	uptime	im. eff.	sp. eff.	im. ops.	sp. ops.	end game	im. hrs.	sp. hrs.	sq. deg.	No. plts
2000 Q2											0.0	0	0.0	0
Q3	130	0.2	0.8	0.5	0.9	0.86	0.65	0.5	0.25	1	5.0	7.6	94.3	10
Q4	185	0.2	0.8	0.6	0.9	0.86	0.65	0.75	0.5	1	12.9	26.0	241.6	35
2001 Q1											0.0	0.0	0.0	0
Q2											0.0	0.0	0.0	0
Q3	130	0.2	0.8	0.5	0.9	0.86	0.65	0.9	0.9	1	9.1	27.4	169.8	37
Q4	92	0.2	0.8	0.6	0.9	0.86	0.65	0.9	0.9	1	7.7	23.3	144.2	31
2002 Q1											0.0	0.0	0.0	0
Q2											0.0	0.0	0.0	0
Q3	0	0.2	0.8	0.5	0.9	0.86	0.65	0.9	0.9	1	0.0	0.0	0.0	0
Q4	0	0.2	0.8	0.6	0.9	0.86	0.65	0.9	0.9	1	0.0	0.0	0.0	0
2003 Q1											0.0	0.0	0.0	0
Q2											0.0	0.0	0.0	0
Q3	0	0.2	0.8	0.5	0.9	0.86	0.65	0.9	0.9	1	0.0	0.0	0.0	0
Q4	0	0.2	0.8	0.6	0.9	0.86	0.65	0.9	0.9	1	0.0	0.0	0.0	0
2004 Q1											0.0	0.0	0.0	0
Q2											0.0	0.0	0.0	0
Q3	0	0.2	0.8	0.5	0.9	0.86	0.65	0.9	0.9	1	0.0	0.0	0.0	0
Q4	0	0.2	0.8	0.6	0.9	0.86	0.65	0.9	0.9	1	0.0	0.0	0.0	0
2005 Q1											0.0	0.0	0.0	0
Q2											0.0	0.0	0.0	0
TOTAL	537										34.7	84.2	650.0	112

Equatorial Survey

The equatorial stripe is 108 degrees long and thus has a footprint of 270 square degrees. The defining characteristic of the equatorial survey is that the stripe will be scanned multiple times during the course of five years, and there will be opportunities for adopting a different spectroscopic strategy to provide information complementary to the other surveys. When the multiple scans are co-added, we will achieve a correspondingly deeper image of the sky. Moreover, the time series enables a unique ability to detect objects which vary in brightness or which change position.

To formulate the baseline for the equatorial survey, the "dark hours" column in Table 3 is adjusted for the time allocated to the outrigger survey. The allocation between imaging and spectroscopy is 0.4 / 0.6 to reflect a deliberate emphasis on imaging. All of the other factors are the same as for the other surveys. This model yields a projection that by the end of the five-year survey, in the equatorial survey we will have imaged 5220 square degrees. This is a repeat factor of 5220 / 270 = 19. In other words, we expect to scan the southern equatorial stripe 19 times in five years.

Table 3. SDSS 5-Year Model Projection – South Equatorial Survey

	dark hrs.	im. frac.	sp. frac.	weath	uptime	im. eff.	sp. eff.	im. ops.	sp. ops.	end game	im. hrs.	sp. hrs.	sq. deg.	No. plts
2000 Q2											0.0	0	0.0	0
Q3	65	0.4	0.6	0.5	0.9	0.86	0.65	0.5	0.25	1	5.0	2.9	94.3	4
Q4	92	0.4	0.6	0.6	0.9	0.86	0.65	0.75	0.5	1	12.8	9.7	240.3	13
2001 Q1											0.0	0.0	0.0	0
Q2											0.0	0.0	0.0	0
Q3	65	0.4	0.6	0.5	0.9	0.86	0.65	0.9	0.9	1	9.1	10.3	169.8	14
Q4	185	0.4	0.6	0.6	0.9	0.86	0.65	0.9	0.9	1	30.9	35.1	579.9	47
2002 Q1											0.0	0.0	0.0	0
Q2											0.0	0.0	0.0	0
Q3	195	0.4	0.6	0.5	0.9	0.86	0.65	0.9	0.9	1	27.2	30.8	509.4	41
Q4	277	0.4	0.6	0.6	0.9	0.86	0.65	0.9	0.9	1	46.3	52.5	868.3	70
2003 Q1											0.0	0.0	0.0	0
Q2											0.0	0.0	0.0	0
Q3	195	0.4	0.6	0.5	0.9	0.86	0.65	0.9	0.9	1	27.2	30.8	509.4	41
Q4	277	0.4	0.6	0.6	0.9	0.86	0.65	0.9	0.9	1	46.3	52.5	868.3	70
2004 Q1											0.0	0.0	0.0	0
Q2											0.0	0.0	0.0	0
Q3	195	0.4	0.6	0.5	0.9	0.86	0.65	0.9	0.9	1	27.2	30.8	509.4	41
Q4	277	0.4	0.6	0.6	0.9	0.86	0.65	0.9	0.9	1	46.3	52.5	868.3	70
2005 Q1											0.0	0.0	0.0	0
Q2											0.0	0.0	0.0	0
TOTAL	1823										278.3	307.8	5217.5	410

Note that the projection for the spectroscopic survey in the equatorial stripe is currently unrealistic: 410 plates obtained, but it takes only 54 plates to cover the area. A scenario where the exposure times are increased to 2 hours from 45 minutes, and where with the increased depth we target each tile three times to sample the larger surface density of objects, succeeds in covering the area to that increased depth in five years. The specific strategy remains to be determined, however. Up to this point, the spectroscopy in the equatorial stripe has been done in the same way as for the North and for the outrigger stripes.