https://doi.org/10.71929/rubin/2584084
Draft version August 19, 2025
Typeset using LaTeX **modern** style in AASTeX7

Survey Strategy and Cadence Choices for the Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST)

R. Lynne Jones,[1] Peter Yoachim,[1] Željko Ivezić,[2,1] Eric H. Neilsen, Jr.,[3] and Tiago Ribeiro[2]

[1] *University of Washington, Dept. of Astronomy, Box 351580, Seattle, WA 98195, USA*
[2] *Vera C. Rubin Observatory Project Office, 950 N. Cherry Ave., Tucson, AZ 85719, USA*
[3] *Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA*

(Dated: 2021-03-11)

ABSTRACT

A summary of survey strategy and cadence choices, simulated and evaluated by the Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST) Scheduler Team, prepared for the Survey Cadence and Optimization Committee (SCOC).

A large telescope survey, covering the entire visible sky repeatedly every few days in multiple bandpasses over the course of ten years, is the core idea of the LSST. An area of about 20,000 square degrees observed under a wide range of conditions to deep coadded limiting magnitudes in bandpasses *ugrizy* enables cosmological studies and studies of the Milky Way structure with unprecedented precision; the same survey, when cadenced well, can serve to open new windows into our understanding of transient and variable stars, and extend our knowledge of small bodies throughout the Solar System by orders of magnitude. The outlines of, and some basic necessary requirements for these goals are outlined in the LSST Science Requirements Document (SRD)[a)]. Finding options for the observing strategy to meet more detailed needs of an even wider range of science goals, as well as building the LSST Scheduler and Metrics Analysis Framework, has been the work of the LSST Scheduler Team with support and input from the astronomical community, including the Community Observing Strategy Evaluation Paper (COSEP)[b)], the Call for White Papers[c)], numerous performance metrics, and guidance from the LSST Science Advisory Committee in their Recommendations for Operations Simulator Experiments[d)]. This report will help enable the SCOC to make its recommendation for the preferred survey strategy.

[a)] ls.st/srd
[b)] https://github.com/LSSTScienceCollaborations/ObservingStrategy
[c)] Document-28382 (Ž. Ivezić et al. 2018)
[d)] Document-32816 (LSST Science Advisory Committee 2019)

Contents

1. INTRODUCTION

Vera C. Rubin Observatory (Rubin) will carry out the Legacy Survey of Space and Time (LSST) over the first ten years of its lifetime (Ž. Ivezić et al. 2019). The LSST is intended to meet four core science goals:

- constraining dark energy and dark matter
- taking an inventory of the Solar System
- exploring the transient optical sky, and
- mapping the Milky Way.

The basic requirements for these goals are described in the LSST Science Requirements Document (SRD; Ž. Ivezić & The LSST Science Collaboration 2018). In practice, the SRD intentionally places minimal quantitative constraints on the observing strategy, primarily requiring:

- A footprint for the 'main survey' of at least 18,000 deg^2, which must be uniformly covered to a median of 825 30-second visits per 9.6 deg^2 field, summed over all six filters, *ugrizy* (see SRD Tables 22 and 23). This places a minimum constraint on the time required to complete the main survey. Simulated surveys indicate that the main survey typically requires 80–90% of the available time (10 years) to reach this benchmark; even with scheduling improvements, it is unlikely that the goals of the main survey could be met with a time allocation significantly below 80%.
- Parallax and proper motion 1σ accuracies of 3 mas and 1 mas/yr per coordinate at $r = 24$, respectively, in the main survey (see SRD Table 26), which places a weak constraint on how visits are distributed throughout the lifetime of the survey and throughout a season.
- Rapid revisits (40 seconds to 30 minutes) must be acquired over an area of at least 2000 deg^2 (see SRD table 25) for very fast transient discovery; this requirement can usually be satisfied via simple field overlaps when surveying contiguous areas of sky.

This leaves significant flexibility in the detailed cadence of observations within the main survey footprint, including the distribution of visits within a year (or between seasons), the distribution between filters and the definition of a 'visit' itself. Furthermore, these constraints apply to the main survey; the use of the remaining time (i.e., in mini surveys) is not constrained by the SRD.

In order to maximize the overall science impact of the LSST, in 2018 the project issued a call for white papers (Ž. Ivezić et al. 2018) requesting survey strategy input. The 46 submitted white papers represent a wide swath of the astronomical community, and work together with the Community Observing Strategy Evaluation Paper

(COSEP)[4] to shape the next stage of the survey strategy evaluation. The contents of these white papers were distilled into several areas for investigation by the LSST Science Advisory Council (SAC) in their advisory response (LSST Science Advisory Committee 2019) to the project.

This survey strategy optimization work is starting from an existing candidate baseline strategy, driven by the basic science goals. A brief introduction to the baseline survey strategy, expanded background of the primary LSST science goals, and concise descriptions of how these goals drive the basic survey strategy and data processing requirements are provided in the LSST Overview paper (Ž. Ivezić et al. 2019). A reference survey simulation (`baseline2018a`), generated by an earlier version of the LSST survey simulation tools (see Section 2), provided an implemented example of this strategy. This starting point for the survey strategy can be described extremely briefly as follows:

- The **main "wide-fast-deep" (WFD) survey**, which covers ∼18,000 deg^2 of sky within the equatorial declination range $-62° < \delta < +2°$, and excluding the central portion of the Galactic plane. Within the main survey, two visits[5] per 9.6 deg^2 field (in either the same or different filters) are acquired in each night, to allow identification of moving objects and rapidly varying transients, and to improve the reliability of the alert stream. These pairs of visits are repeated every three to four nights throughout the period the field is visible in each year (other nights are used to maximize the sky coverage). Each field in the main survey receives about 825 visits throughout the ten years of the LSST, spread over the six LSST filters *ugrizy*. The quantitative SRD constraints on area coverage, number of visits, parallax and proper motion errors, and rapid-revisit rate (40 seconds – 30 minutes) apply to visits obtained in the main survey.

- The set of five **Deep Drilling Field candidate mini surveys**, consisting of five specific field pointings for a total of ∼ 50 deg^2, which are observed with a much denser sampling rate. These mini surveys use a similar sequence of visits; the fields are observed every three to four days, but in a sequence of multiple *grizy* exposures during gray and bright time, and then multiple sequential *u* band exposures during dark time. The current deep drilling mini survey fields are aimed at extragalactic science, providing a 'gold sample' to calibrate the main survey, and to discover Type Ia supernovae.

- The **Galactic Plane candidate mini survey** covers the central portion of the Galactic plane that is not included in the main survey, centered around $|l| = 0°$ and covering ∼ 1860 deg^2. It is observed at a much reduced rate compared to the main survey, with about five times smaller total number of observations per

[4] The github repository containing the living source for the COSEP is https://github.com/LSSTScienceCollaborations/ObservingStrategy

[5] A 'visit' here is an LSST default visit, which consists of two back-to-back 15 sec exposures, for a total of 30 sec of on-sky exposure time. These back-to-back exposures are always in the same filter, separated only by the 2 second readout time.

field (30 visits per field and per filter, in *ugrizy*), so as to provide astrometry and photometry of stars toward the Galactic center but without reaching the confusion limit in the coadded images. There is no requirement for pairs of visits in each night in this area.

- The **North Ecliptic Spur candidate mini survey** covers the area north of $\delta = +2°$ to 10° north of the Ecliptic plane and is intended to observe the entire Ecliptic plane for the purpose of inventorying the minor bodies in the Solar System. This area ($\sim$ 4160 deg^2) is observed on a schedule similar to the main survey, although with a smaller total number of visits per field and only in filters *griz*.
- The **South Celestial Pole candidate mini survey** covers the region south of the main survey, to the South Celestial Pole, $\sim$ 2315 deg^2, including the Magellanic Clouds. This mini survey is observed with a strategy similar to the Galactic Plane mini survey, with 30 visits per field per filter in *ugrizy*, and without requiring pairs of visits. This provides coverage of the Magellanic clouds, but without committing extensive time as these fields are at high airmasses from the LSST site.

This report covers the LSST Survey Strategy team's experiments with the LSST scheduler to address the optimization questions raised by the SAC. These questions include:

- How should the WFD footprint be defined?
- What should the cadence of visits within the WFD look like? This includes both the intra-night cadence and the inter-night cadence throughout the season.
- What is the impact of varying the footprint for mini-surveys?
- Can we leverage twilight observing?
- How should the Deep Drilling fields be distributed and what cadence should be used for their observations?
- What are the impact of ToO proposals, particularly gravitational wave followup?

These questions are aimed at ensuring the best possible science return from the LSST.

2. SURVEY SIMULATOR OVERVIEW

In operations, the LSST needs an automated scheduler to appropriately plan and execute about 1000 visits per night. Prior to operations, we have a need to use the same scheduler to understand the range of possible survey strategies and their science potential; even in operations it is useful to run the scheduler in a 'simulation' mode in order to evaluate the future impact of changes in observatory hardware or changes to the observing strategy. As such, we need a robust scheduler, together with high-fidelity model inputs for the telescope operations and observing telemetry.

2.1. *The Model Observatory*

2.1.1. *Telescope Model*

The physical telescope operations are modeled using the LSST software package `ts_observatory_model`. This package includes a kinematic model of the telescope, with appropriate acceleration/deceleration and maximum velocity limits, including requirements for sequencing (changing the filter before slewing, for example). It also enforces requirements needed before image acquisition, such as the settle time after slewing and the active optics open- and closed-loop acquisition times. Other important considerations are the extent of cable wrap due to azimuth slews or camera rotation. The parameters for the telescope model are configurable, coming from the Telescope and Site and Camera teams. These parameters are largely unchanged from F. Delgado et al. (2014); a subset of these parameters are described in Table 1.

2.1.2. *Cloud Model*

The cloud model is based on historical cloud sky coverage data from Cerro-Tololo Inter-American Observatory (CTIO), from the ten year period 1996 to 2005.

Min altitude	20 deg
Max altitude	86.5 deg
Camera readout	2 sec
Shutter time	1 sec
Filter change time	120 sec
Number filters mounted	5
Azimuth slew settle time	1 sec
Closed Optics Loop Delay	36 sec (when > 9 deg altitude change)
Approximate azimuth slew time	$t_{slew\,Az} = 0.66\,\mathrm{sec/deg} * \delta Az(\mathrm{deg}) + C^{Az}$ $C^{Az} = -2$ s; $t_{slew\,Az}$ min $= 3$ sec
Approximate altitude slew time	$t_{slew\,Alt} = 0.57\,\mathrm{sec/deg} * \delta Alt(\mathrm{deg}) + C^{Alt}$ $C^{Alt} = 3$ s for slews below 9°, $C^{Alt} = 37$ s for slews above 9° alt

Table 1. A subset of survey strategy relevant ts_observatory_model parameters and slew time approximations.

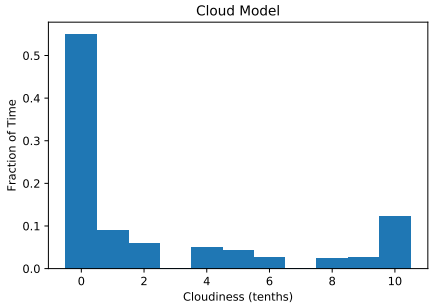

Figure 1. The distribution of cloudiness as measured at CTIO. We model the observatory as closed for cloud levels above 3/10.

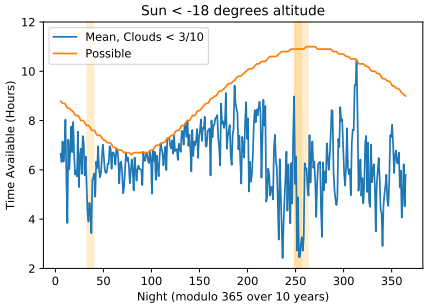

Figure 2. The total amount of possible time per night and the average amount of time after removing weather downtime. Shaded regions show scheduled downtime. Our modeled unscheduled downtime is not included in the plot. The year progresses from left to right in the plot starting on October 1st. The variable length June shutdown starts around Night 250. This general weather and maintenance schedule is the same for all simulations.

The SOAR telescope reports losing 15.3-33.4% (mean=22.9%) of science time to weather from 2014-2018[6]. This is consistent with the weather downtime reported by Gemini South (private communication).

If we model the observatory as closed when the sky is 30% cloudy or cloudier, we reach a weather downtime of 29.8%. While we expect some observations will be possible in 30% cloudy skies, this cutoff also accounts for other weather related closures (humidity, wind, dust, etc).

2.1.3. *Seeing Model*

Simulations completed starting in 2020 use a revised database for the atmospheric seeing. The revised database, like its predecessor, is based on seeing measurements from the Geminin South DIMM, located at the same site as Rubin Observatory. We derived predicted delivered image FWHM at 500 nm at zenith from the reported DIMM measurements using the approximation of the von Kármán turbulance model given in Tokovinin (2002) and an outer scale of 30 meters, and validated this relationship between DIMM measurements and seeing by comparing these derived values

[6] http://www.ctio.noao.edu/soar/content/soar-observing-statistics

to the image quality measured from the Gemini South GMOS instrument. We also tested the DIMM data by deriving a seeing value and comparing the result to the seeing measured by the DECam imager on the Blanco telescope at CTIO, a few miles away.

For most time samples in the simulation database, we generated seeing data by resampling seeing derived from the DIMM into 5 minute intervals, and shifting it ahead 4748 days (13.000 tropical years). For example, the seeing for 2022-01-01 in the simulation database is taken from the DIMM seeing on 2009-01-01. Thus, most of the ten simulated years use seeing values that replay ten historical years.

There is, however, significant time for which no DIMM data is available, for example due to clouds or equipment failure. We used a model of $log(r_0)$ (where r_0 is the Fried parameter) derived from the DIMM data to generate artificial seeing values for these times. This model has several components:

- a yearly sinusoidal variation in $log(r_0)$ to include seasonal variation,
- a smooth (years timescale) fit to the residuals with respect to the seasonal variation to represent multi-year trends in seeing,
- a 1st-order autoregressive series (damped random walk) to represent variations in the nightly seeing, and
- another 1st-order AR series to represent variations on a 5-minute timescale within a night.

Artificial data generated according to this model therefore maintains the night to night and short term distributions and correlations present in the DIMM data, and follows seasonal variations and longer term trends in the DIMM data surrounding it.

The FWHM in any particular image of a simulation is calculated from the atmospheric contribution (FWHM at 500 nm at zenith from the database) and the telescope system contribution (based on engineering specifications), correcting for wavelength and airmass.

2.1.4. *Sky Brightness Model*

The observatory model includes a model for the sky brightness. The model is built mostly from the ESO sky brightness model which includes upper and lower atmosphere emission lines, airglow continuum, scattered lunar light, and zodiacal light. In addition, we have added a twilight model fit from all-sky camera observations at the site. The sky brightness model does not include human generated light pollution. While the ESO model does include the ability to scale the airglow component with solar activity, we use the default mean solar activity throughout. Compared to all sky camera observations, the sky brightness model has RMS deviations of ∼0.2-0.3 magnitudes per square arcsecond (P. Yoachim et al. 2016).

With so many independent components, the sky brightness is potentially the most computationally expensive aspects of the simulations. We pre-compute sky brightness

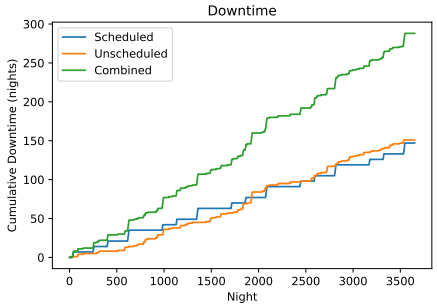

Figure 3. The simulated scheduled and unscheduled downtimes over 10 years.

maps in all six Rubin filters in 5-15 minute time steps, depending on how rapidly the sky brightness is changing (twilight requires finer time steps than the middle of the night) which can then be rapidly interpolated to exact times.

2.1.5. *Maintenance Downtime Model*

The observatory model includes both scheduled and unscheduled downtime. Figure 3 shows we simulate approximately 10% of time lost to maintenance. The scheduled downtime allowance is currently about 22 weeks over the full 10 year survey. This is taken in either two one week periods twice a year, or a single two week period in alternating years. The unscheduled downtime allowance is approximately 21 weeks, in variable amounts of time, often as short as a single night. The scheduled downtime is planned during the same periods that are most likely to be bad weather, when possible. In the future, scheduled downtime may be shifted within a month to better align with the full moon.

2.2. *The Scheduler*

Optimally scheduling telescopic observations is a traditionally difficult problem. Most observatories have typically scheduled observations by hand. The Las Cumbres Observatory (LCO) and Zwicky Transient Factory (ZTF) have implemented Integer Linear Programming techniques to optimize their scheduling (S. Lampoudi et al. 2015; E. C. Bellm et al. 2019). With integer programming, potential observing time is quantized into blocks and an optimization algorithm is used to maximize a user-defined objective function. Integer programming is difficult to use for Rubin because we have multiple science goals which are intended to be serviced simultaneously. Thus, there is no easily-defined function which can be maximized when scheduling Rubin. D. Rothchild et al. (2019) simulated Rubin observations with a very fast deterministic scheduler, essentially repeating a fixed raster pattern mostly along the meridian. This algorithm showed great promise, but had several downsides (such as occasionally pointing at the moon). For the Rubin scheduler, we follow the example in E. Naghib et al. (2019) and use a Markov Decision Process (MDP) to select most of the observations. With a MDP, observing decisions are made in real-time based on the current conditions and previously completed observations. The MDP relies

on the construction of basis functions which encapsulate the different objectives and constraints of the problem. The advantage of the MDP is that well-constructed basis functions can result in the scheduler having few free parameters which are easily optimized. The disadvantage is that writing good basis functions often require the author have extensive domain-specific knowledge of the problem. For example, while astronomers say things like "try to take observations at low airmass", this by itself does not make a good basis function because there can be declinations in a survey footprint which never reach low airmass.

The Rubin scheduler is designed to provide real-time decisions on where and how to observe. Because we expect there to be weather interruptions, we need a system that can recover quickly. Unlike other traditional telescope schedulers, we do not try to optimize a large number of observations in advance, but rather use a decision tree along with a modified Markov Decision Process. The scheduler behavior is set by a small number of free parameters that can be tuned.

Our baseline scheduler uses a three tier decision tree when deciding what observations to attempt.

2.2.1. *Tier 1: Deep Drilling Fields*

The first tier of the decision tree is to check if there are any deep drilling fields that should be executed. We typically have five DDFs in a simulation.

For a DDF to be eligible to send a sequence to the observing queue, it must

- Not currently be twilight (-18 degrees for DDFs)
- Have enough time to finish a sequence before twilight begins
- Be in its target hour angle range
- The moon must be down (DDFs typically include visits in multiple filters, including some or all of u, g, r or i).
- The DDF must not have exceeded its limit of observations (typically $\sim$1% of the total number of visits)

If the DDF has not fallen behind (fallen below its desired fraction of survey visits by some threshold), it will space sequences by at least 1.5 days. There is also a check to see if the DDF will be feasible and better observed later in the night, in which case no observations are requested.

If the above conditions are met, the DDF sends it's sequence of observations to the queue to be executed. There are currently no attempts at recovery if a sequence is interrupted.

The spatial position of the DDF is dithered nightly up to 0.7 degrees. The camera rotator is also varied nightly to be between -75 and 75 degrees with respect to the telescope. The optimal dithering patterns for the DDFs are yet to be determined.

Name	RA (Deg)	Dec (Deg)
ELAISS1	9.450	-44.000
XMM-LSS	35.708	-4.750
ECDFS	53.125	-28.100
COSMOS	150.100	2.182
EDFS	58.970	-49.280
EDFS	63.600	-47.600

Table 2. The location of the deep drilling fields used in our simulations.

2.2.2. *Tier 2: The Blobs*

If there are no DDFs requesting observations, the decision tree moves to the second tier. This tier is the survey workhorse, executing ∼80% of the simulation visits. This tier will only request observations if it is not currently twilight, and there is at least 30 minutes before twilight begins. The name 'blob' refers to the fact that this tier selects groups of fields for visits that are spatially close to one another; thus a 'blob'.

A modified Markov Decision Process (MDP) is used to decide what sky region (blob) and filter combination to observe given the current conditions and observation history. Briefly, the MDP balances the desire to observe areas 1) that are closest to the optimal possible in terms of 5-sigma depth, 2) which have fallen behind the specified desired survey footprint (have obtained a smaller fraction of the overall survey visits), 3) are near the current telescope pointing to minimize slew time, and 4) in the currently loaded filter to minimize filter changes. In addition to these core components, the MDP includes a mask around zenith, a 30 degree mask around the moon, and small masks around the bright planets (Venus, Mars, Jupiter). The end product of the MDP is a reward function that ranks the desirability of every point in the sky. Because this tier does not execute in twilight, we assume the reward function is relatively stable on 40 minute timescales.

A sky area around the reward function maximum that will take ∼22 minutes to observe (∼35 pointings - a blob) is then selected. If possible, the area is selected to be be contiguous. The exact position of the telescope pointings are determined by the sky tessellation, which is randomly oriented for each night to provide a spatial dithering between nights. The camera rotator angle (relative to the telescope) is also randomized between ±80 degrees each night.

A traveling salesman algorithm is used to put the pointings in an order that minimizes the slew time. The list of pointings are then repeated, usually in a different filter, ensuring moving objects can be detected. One of seven possible filter combinations is used: $u+g$, $u+r$, $g+r$, $r+i$, $i+z$, $z+y$, or $y+y$. We use 30 second visits for the majority of simulations. The official baseline uses visits comprised of two 15 second snaps.

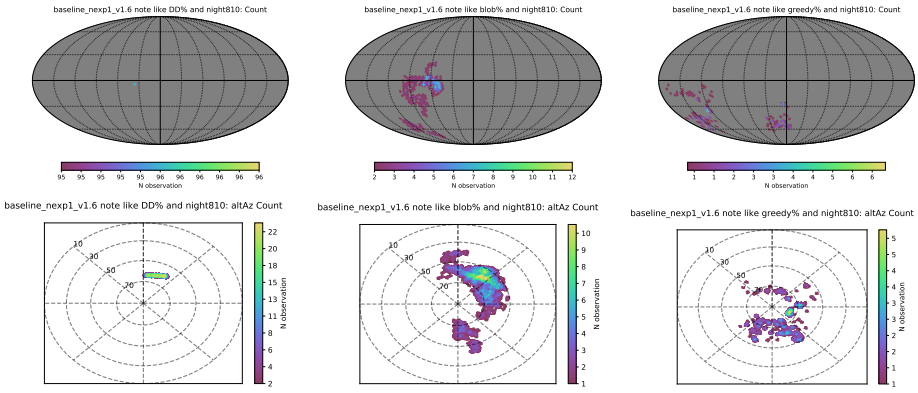

Figure 4. Examples of how the three scheduler tiers execute during a single night. Left panels show how a DDF sequence was observed during the night. Middle panels show observations taken as part of blob pairs. Right panels show the greedy observations taken in twilight time. The panels from left to right show the different decision tiers the scheduler uses, with the DDFs as the top tier and the greedy algorithm as the bottom tier. The top row of panels show visits in RA/Dec coordinates, the bottom row show the same visits in Alt/Az coordinates.

2.2.3. *Tier 3: Greedy*

If it is during morning or evening twilight, or close to morning twilight, the DDFs and Blob surveys will pass and the decision tree goes to the third and final tier, the greedy surveys.

The greedy surveys use a similar Markov Decision Process as in Tier 2, but rather than selecting large areas of sky to observe, the survey selects a single pointing at a time. No attempt is made to observe greedy scheduled observations in pairs of visits. The twilight period is short and the conditions rapidly changing, so pairs may not be optimal; however this is an area for potential improvement. Since this tier is primarily used in twilight time, it only schedules observations in the redder filters r, i, z, and y.

As with the Blob tier, the sky tessellation orientation is randomized each night so the final survey is spatially dithered.

2.3. *Filter Mounting Schedule*

In addition to the observations scheduler, we have a separate scheduler that decides which five filters should be loaded for the start of each night. By default, we mount redder filters (*grizy*) when the moon is more than 40% illuminated and bluer filters (*ugriy*) closer to new moon. (See Section 4.1 for more information on the choice of when to swap the filter.)

3. METRICS

There are many options for evaluating the output of the survey strategy experiments, including high-level science-oriented metrics and more basic metrics measuring simple changes in survey characteristics. One of the primary goals for the LSST Metrics Analysis Framework (MAF) package was to make it easier for both the project and community members to write metrics to evaluate these outputs. This effort has had some significant successes; SRD-level metrics have been written that cover the primary requirements for the SRD, the DESC working groups have made good progress in writing metrics for their evaluation of the simulations, and the Solar System collaboration has contributed substantial metrics. In other areas, it has been more difficult for the community to engage and contribute directly to MAF; for some of these areas, we have been able to help get metrics running, but clearly there are areas which are lacking definitive metrics. Many of the areas which are lacking relate directly to time domain studies, a critical area for the LSST (we do have simple periodic variable detection and period determination metrics, as well as SNIa, TDE and microlensing metrics, and the ability to quickly generalize these metrics for other kinds of transient lightcurves). We acknowledge this issue and look forward to working with the community to address this, and the larger concern of potentially important, but missing, metrics.

Here we make a brief summary of some of the top-level science-related metrics. There are thousands of metrics which are run as part of standard MAF analysis (many of which relate to simple analysis of observation metadata like seeing, airmass, sky brightness, etc.); for broad comparisons between simulations we pick a very limited subset of these metrics intended to discover or highlight differences between the simulation survey strategies or to cover major areas of science. All of these metrics are available in github repos in either `sims_maf` or sims_maf_contrib. We also describe the 'radar plots' (the 2-d plots holding representations of these metric results across multiple simulations) at the end of this section.

3.1. *SRD Metrics*

The SRD metrics are designed to cover the primary science requirements laid out in the SRD; the most relevant of these relate to the number of visits per pointing across the WFD region, the area of the WFD region, the parallax and proper motion errors and the number of rapid revisits (on timescales between a 40 seconds to 30 minutes) per point on the sky. While we check all of these metrics for all runs, the most sensitive to changes in the survey strategy is the number of visits across the WFD, tracked in the fO metric, since we are often attempting to distribute visits into other parts of the sky for other science.

The fO metric calculates the total number of visits per point on the sky, then calculates how much area is covered with how many visits. This can be summarized across two axes; the amount of area that receives at least 825 visits per pointing ('fO

Area') or the median (or minimum) number of visits that the most frequently visited 18,000 square degrees receives ('fONv MedianNvisits' or 'fONv MinNvisits'). The first version, fO Area, tends to be somewhat unstable; the survey hardly ever observes more than 18,000 square degrees to at least 825 visits, because we don't program in larger WFD areas, but if the number of visits across the WFD area falls below 825, the resulting fO Area value will fall rapidly (because we cover the sky uniformly). While fO Area is useful to check, a more useful number is fONv MedianNvisits or MinNvisits. The value of fONv MedianNvisits tells us how many visits the typical field in the top 18,000 square degrees receives; fONv MinimumNvisits tells us the fewest number of visits any of those top 18,000 square degrees received. Typically we see fONv MedianNvisits scales more smoothly with the fraction of visits devoted to WFD and likely represents science metrics that depend on having a reasonably large amount of visits over the entire WFD well.

The radar plots use fONv MedianNvisits, the Median Parallax Error, and the Median Proper Motion Error. The astrometry metrics both assume an r=20 magnitude star with a flat SED. When plotted in radar plots, we compare the reciprical of the astrometry uncertainties so that larger values on the radar plot can always be interpreted as "better".

3.2. *Solar System Science Metrics*

Solar System science metrics include discovery metrics (with various discovery criteria, such as detections in 3 nights with pairs of visits within a 15 night window) and characterization metrics (ie. how many colors for objects can we measure, and can we determine a light curve or even shape measurement from the lightcurve), contributed by both project and science collaboration. The most important metric for solar system objects is discovery; finding the objects is the first priority. Characterization metrics are secondary metrics. For each of these metrics, we generate input observations using an appropriate solar system population: Potentially Hazardous Asteroids (PHAs) and Near Earth Objects (NEOs) based on a model by M. Granvik et al. (2018), Main Belt Asteroids (MBAs) and Jovian Trojans based on the S3M model from T. Grav et al. (2011), and TransNeptunian Objects (TNOs) based on the L7 model from the Canada France Ecliptic Plane Survey (CFEPS) (J. J. Kavelaars et al. 2009; J. M. Petit et al. 2011). These populations move at varying rates and cover varying amounts of the sky. NEOs move over much of the sky during the lifetime of the survey, so are less sensitive to footprint variations, but tend to have much more strongly varying brightnesses, thus are sensitive to the number and timing of visits (must be observed when they are bright). TNOs move very slowly, not more than a few fields of view over the lifetime of the survey, so are quite sensitive to footprint, however they are relatively consistent in their brightness; thus they are less sensitive to the overall number of visits at a particular point in the sky, once a threshold has been met.

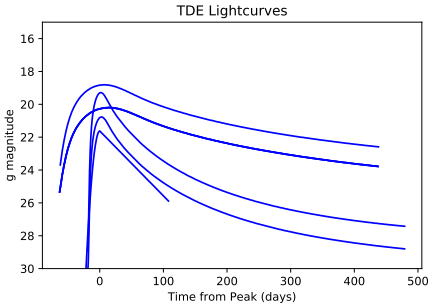

Figure 5. Simulated TDE lightcurve shapes.

For each of these populations, we calculate the population completeness due to discovery with the LSST at the end of 10 years (not including previous surveys) with the currently Moving Object Pipeline baseline criteria; 3 nights with pairs of visits within 15 nights at a range of absolute magnitude H (approximately the size of the object) and then take the completeness at an H value near peak completeness and an H value that is relatively close to 50% completeness in the baseline; these completeness values are the summary metrics we track across various runs to compare them here.

The radar plots use the completeness for bright (H=16) NEOs, faint (H=22) NEOs, and bright (H=4) TNOs.

3.3. *Number of Stars*

We use a simulated MW stellar catalog from Galfast (M. Jurić et al. 2008; M. Jurić 2018) along with the survey coadded depth to estimate the number of stars that would be detected at the 5σ level in i. Comparison with the TRILEGAL galaxy model (L. Girardi et al. 2005, 2012) gives similar results.

The number of stars is primarily sensitive to the footprint definition, and decreases dramatically (by about a factor of 2) when there is no coverage of the galactic plane. An extended survey footprint (such as increased coverage toward the north) also increases the number of stars. Because we are only computing the metric in the i filter, the metric is also sensitive to the depth in i, and thus we can see some variation if, e.g., a simulation pushes more i observations to twilight time.

The radar plots use the total number of stars over the entire footprint down to the coadded limiting magnitude. We do not include a crowding correction.

3.4. *Tidal Disruption Events (TDE)*

We use TDE lightcurves from the community to generate a sample of TDE events distributed uniformly on the sky and uniformly over the 10 year survey. Figure 5 shows the lightcurve shapes. When analyzing a detected light curve, we test three crtiera

- If it is detected twice pre-peak in any filters

- If there is one detection pre-peak, and detections in at least 3 filters within 10 days of peak

- If there is one detection pre-peak, one detection in u and any other band near peak, and u plus any other filter post-peak.

When requiring both a color in any filter and u band measurements during the TDE event, this metric is exceedingly sensitive to the number and cadence of u band visits, with the number of detected TDEs scaling linearly with the number of u band visits and preferring visits spread more uniformly over time. In other configurations, when requiring observations pre-peak or just a color in any filters, it is primarily sensitive to the frequency of observations and whether pairs are obtained in the same or mixed filters.

The radar plot uses the TDE some color plus u band metric output.

3.5. *Fast Microlensing*

We use microlensing light curves contributed from the community. For all the events, we assume an r=22 magnitude star with a flat SED is being magnified.

We calculate both a Fast (crossing times of 1-10 days) and Slow (crossing times 100-1,500 days) microlensing metric. They are distributed on the sky proportionally to stellar density squared as measured from TRILEGAL galaxy model. Due to this spatial distribution, both the Slow and Fast microlensing metrics are primarily sensitive to survey footprint. Footprints without galactic plane coverage cut the number of detected microlenses by approximately 75% while footprints with heavy galactic plane coverage can increase the number of microlenses by a factor of 2 or more. The Fast microlensing metric is also sensitive to the number and cadence of u band visits, preferring u band visits spread more uniformly over time.

The radar plot uses the Fast microlensing metric. We find the slow microlensing events are so slow they are detected at a very high rate regardless of survey strategy.

3.6. *Number of Galaxies*

The estimated expected number of galaxies, across the entire survey footprint, is calculated using GalaxyCountsMetric_extended, from sims_maf_contrib. The number of galaxies is estimated based on the coadded depth using redshift-bin-specific power-laws, based on mock catalogs; the construction of these power laws and normalization is described in H. Awan et al. (2016). The overall number of galaxies tends to increase with increased depth, and more so when more of the survey footprint is distributed in lower dust extinction areas. The number of galaxies also increases when the survey filter distribution is redder, rather than bluer.

The radar plots use the total number of galaxies down to the coadded limiting magnitude over the entire survey footprint.

3.7. *DESC WFD Metrics*

The DESC has contributed several metrics evaluating the performance of the WFD for various areas of relevant science. Many of these metrics are built on calculating a subset of the survey footprint that meets the requirements of coverage in all 6 filters, less than a specified level of dust extinction (E(B-V) < 0.2) and greater than a specified coadded depth in i band ($i > 25.9$ at 10 years), calculated using ExgalM5_with_cuts. This represents the extragalactic science footprint.

3.7.1. *Static Science*

Over this extragalactic footprint the following metrics are calculated for general 'static science'.

- Median coadded depth in i band
- Standard deviation of the coadded depth in i band
- The area of the selected footprint
- A 3x2 point Figure of Merit emulator

. These metrics are very sensitive to footprint coverage and depth, as well as desiring uniformity in the coadded depth to minimize corrections during later analysis. The radar plot uses the 3x2point FoM.

3.7.2. *Weak Lensing*

The same footprint is used to calculate the number of visits per point in the footprint (WeakLensingNvisits); this is used as an approximate metric evaluating weak lensing systematics.This metric is sensitive to footprint coverage and depth. The radar plot uses the mean number of visits across the extragalactic footprint.

3.7.3. *Large Scale Structure*

The number of galaxies within this same footprint is used as a metric to approximate large scale structure results (DepthLimitedNumGalaxies), using the same GalaxyCountsMetric_extended as above, but limiting the result to the selected footprint.

3.7.4. *SNe Ia*

We use SNe Ia light curves from the PLAsTiCC challenge. SNe are distributed uniformly on the sky.

For each supernova, we check:

1. Is the supernova detected in any filter?
2. Is there a color detected (detected in 2 filters within 0.5 days)?
3. Is it possible to measure the rise slope (detect an increase of 0.3 mags in a filter pre-peak)?

4. Is the light curve "well sampled" (if the light curve duration is divided into tenths, are there detections in 5 unique bins)?

There are several versions of this metric, using different criteria for observations. We call the supernova "Detected" if it meets criteria 1, it is "Pre-peak" if it meets criteria 2 and 3, and is "Well-sampled" if it meets criteria 4. The simple Detected metric is sensitive to a combination of survey footprint and number of visits, preferring more area as long as a minimum number of visits spaced uniformly over time are available. The Pre-peak metric, which may be more useful for detections before follow-up, is most sensitive to requiring visits to be obtained with mixed filter pairs, with a lesser preference for visits being spaced more evenly over time (such as in non-rolling cadences). The Well-sampled metric is primarily sensitive to the cadence of visits, preferring visits to be spaced evenly over time.

The radar plot uses the metric which demands criteria 2 and 3 from above. Thus, we are mostly measuring how well we are producing SNe alerts that can act as triggers for others to follow up. The DESC group has developed metrics for measuring how well SNe are observed by Rubin alone, and we will be incorporating these into MAF soon.

3.8. *Radar Plots*

To help compare multiple science and SRD metrics across runs, we make use of radar plots, a 2-dimensional plot containing metric results along various, radially-distributed axes. In each radar plot, we typically normalize values to a baseline run and plot the fractional change in metric values in the radial direction. For metrics that are measured over the entire sky (e.g., Parallax, Proper Motion, Weak Lensing), we use the median. For the parallax and proper motion metrics, the inverse of the errors are compared. When there are particularly large changes in metrics, we will generate a pair of radar plots with different radial ranges to make comparisons easier.

In some cases we make radar plots of the median coadded depth in each filter. For coadded depth, we plot magnitude difference in the radial direction, with larger values indicating deeper depths.

Almost all of the metrics in the radar plot show highly statistically significant changes as the survey strategy changes. It is worth remembering that the simulations themselves have some level of uncertainty, as change in the 'real-life' weather or status of the observatory will lead to changes in the observing history and then changes in the scheduler choices to maintain the overall survey strategy guidelines. For some metrics with particularly small measured values, these small differences between runs can result in large metric differences, effectively making the metric results somewhat noisy. The TDE metric is one of these such metric; out of 10,000 simulated TDE lightcurves, the baseline run only observes about 200 of these with visits that meet the 'some color plus u' criteria, thus implying that variations of up to about

7% in the TDE metric value can be expected even if the simulations are statistically similar.

4. SURVEY STRATEGY EXPERIMENTS

The SAC report raised a series of questions and identified suggested simulation experiments to run. This can be categorized as follows:

- Experiments with the WFD footprint (survey footprint variations)
- Experiments with the WFD cadence (note that unless specifically required, we use the same general cadence for the entire sky);
  - Compare individual visits of 2x15s exposure and 1x30s exposure
  - Compare pairs in the same filter vs. different filters, and the effect of triplets of visits
  - Add short (1 or 5 second) visits; test 60 second u band visits.
  - Variable exposure times for uniform depth
- Experiments with rolling cadence, with 2, 3 and 6 declination bands
- Experiments with mini-surveys to the North, South and through the Galactic Plane (essentially, survey footprint variations)
- Experiments with twilight observing
- Experiments with Deep Drilling field cadences
- Tests of Target of Opportunity (ToO) observing

We have explored these areas, along with a few other questions that have arisen over the course of this work, using 'families' of simulations. In these families, we vary a particular parameter of the survey strategy to look for the impact on science. Sometimes the experiments requested by the SAC cross multiple families of these simulations – the general 'WFD cadence' question is addressed with several families investigating different options for cadence variation – and sometimes the impacts to given science goals come from multiple families – the most common being a combination of survey footprint and cadence. Often the impacts are minimal; the baseline LSST survey strategy covers most of the requirements, and these variations are relatively small. Occasionally there are impacts that are much larger, and these are important to note.

The starting point was the existing baseline survey strategy, `baseline2018a`, consisting of the Wide-Fast-Deep (WFD) survey, five Deep Drilling Fields (DDFs,) and Galactic Plane (GP), North Ecliptic Spur (NES) and Southern Celestial Pole (SCP) minisurveys. The existing baseline used 2x15s exposures per visit, and most visits were in the same filter (although this was not enforced). Standard observing started and ended at 12 degree twilight. The general strategy from this simulation was ported to the new scheduler code and (approximately) recreated as the baseline survey strategy.

In the course of working through these simulation experiments, we have issued several releases – sets of simulations which explored parts of the SAC questions using a particular version of the scheduler and simulator code. With each release, we found some improvements or updates to the scheduler or simulation code and also added new simulations investigating new questions. With each release, we typically re-ran the previous set of experiments, although sometimes families of simulations were dropped or modified due to what we learned from the previous release. Release notes can be found on community.lsst.org[7].

The families of simulations relevant for this report come primarily from Feature Based Scheduler (FBS, Section 2.2) release 1.5, 1.4 and 1.6. The FBS 1.5 families are the primary 'experiment' set. One of the FBS 1.4 families was used to set some of the default parameters about u band visit pairing and u band filter load/un-load times; we describe this first. Then we describe the FBS 1.5 families of runs which explore single aspects of survey strategy variations, including the FBS 1.6 experiment families for rolling cadence and filter choice. In the next section we will discuss the FBS 1.6 candidate baselines, where multiple aspects of survey strategy are varied at the same time.

4.1. *FBS 1.4* `u_pairs`*: u Filter Pairing and Filter Load/Unload Time*

One of the early concerns from the SAC was about the u band filter load/un-load time. As the camera can only hold 5 filters at a time, one filter must always be unavailable. We currently swap u band with y band, depending on the phase of the moon. The SAC initially suggested keeping the u band filter in the camera for a very limited time, only a few days around new moon. The driving concern here was to restrict u band usage to the darkest period of the month (increasing the depth in u band) and to allow more consistent sequences of *grizy* for DDFs, as the previous version of the scheduler code would only trigger these sequences when all of the filters were available (and thus would not trigger when u band was in the camera).

As an opposing tension, there was some concern that limiting u band availability could cause problems for classification of transient sources; u band brightness is an important distinguishing feature for many of these objects, particularly Tidal Disruption Events (TDEs). Part of the requirement here is obtaining u band photometry in close proximity to g or r band photometry.

To evaluate all of these issues, we created the `u_pairs` family of simulations. In this family, we only take u band observations paired ($\sim$ 22 minutes later) with g or with r. We vary the timing of when the u band filter is loaded into the camera from 15% to 60% lunar illumination; see Figure 6 for a translation between lunar illumination and days from new moon. We specifically add a variation in the number of u band visits

[7] The Survey Strategy section of community.lsst.org is available at https://community.lsst.org/c/sci/survey-strategy/

per pointing (the weight of the u band footprint) over 1, 2 or 4 times the baseline footprint, in order to more fully explore the impact on transients.

We find that it is necessary to keep the u band in the camera until about 40% lunar illumination to meet SRD requirements on the number of visits per pointing evenly across the sky without increasing the u band weight in the survey footprint. A shorter period of time is long enough to take enough u band visits over the sky *if* the time available is distributed as evenly as the footprint over which those u band visits are required; however, there are seasonal night length and weather variations which make the resulting sky coverage patchy if the number of nights available with u band are too few. See Figure 8. This is primarily a problem with u band, rather than other bandpasses, because of the relatively few number of visits in u band compared to other bands.

These simulations also showed that, no matter how long the u band was available beyond full moon, the basis functions which drive scheduling inside the FBS are able to limit visits in u band to only when the sky brightness in u is low – the moon down, or far away from the field, or a small percent illuminated, and a low airmass for the field. Thus, expanding the period of time the filter is available by itself does not result in lower five sigma limiting depths. See Figure 7.

And finally, we addressed the issue of the *grizy* sequences by adding code that let the DD sequences be more flexible, using whichever filters were available. In all newer simulations, u band is part of the DD sequences instead of separate, and the sequences range over whatever *ugrizy* filters are available at any given time; this has the positive effect of reducing the large gaps between sequential visits in the same filter that were previously a function of lunar phase.

The resulting science trades can be visualized in the radar plots; see Figure 9. The largest change is in the TDE metric, which we can use as a representative of a portion of fast transient science; more u band visits and longer availability results in an increase in the TDE metric results, at a slight cost to the other science cases that don't benefit from additional u band coverage. A closer look at the full TDE metric results show some simple scaling with the total number of u band visits (see Figure 10). One notable point is that the metric results were improved even with a fairly simple change in how the u band visits were acquired; instead of requesting u band visits in singletons, these runs requested u paired with g or r. This improves transient science, although it does increase the amount of time that u band must be available in the camera.

Based on the minor costs to other science, especially after the adjustments made to the scheduler code regarding the DD fields, this family of simulations led us to adjust the baseline survey strategy defaults for all FBS 1.5 runs. For all further runs, we load and unload the u band filter at 40% lunar illumination and pair u band visits with g or r band. We maintain the survey footprint filter ratios at standard.

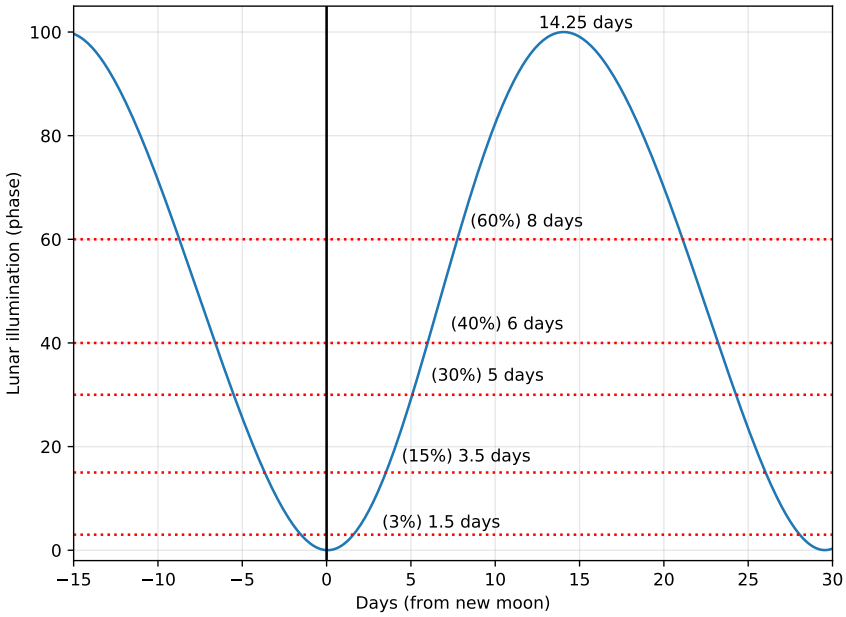

Figure 6. Relationship between lunar illumination (used as the constraint for when to change the available filters) and days from new moon.

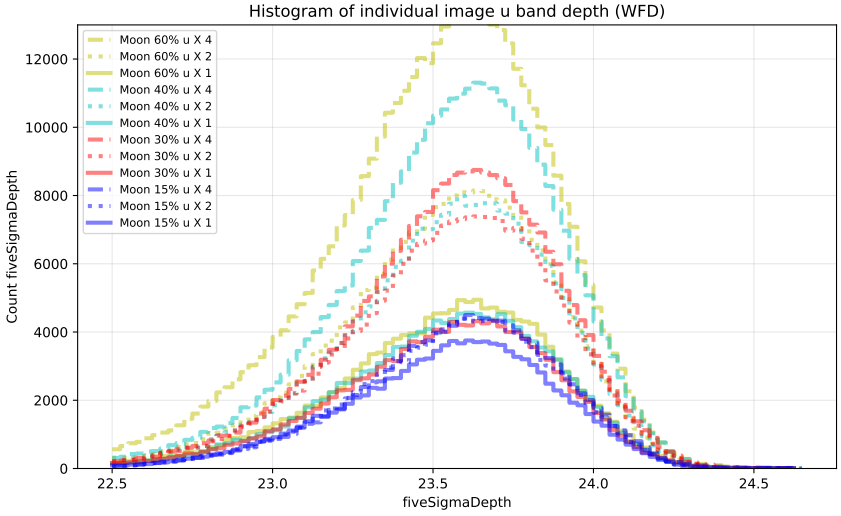

Figure 7. The five sigma depth in u band visits in each simulation in this family. Regardless of the time of the u band filter swap (15, 30, 45 or 60% lunar illumination), when the number of visits fits into the dark time available, the five sigma depths remain comparable.

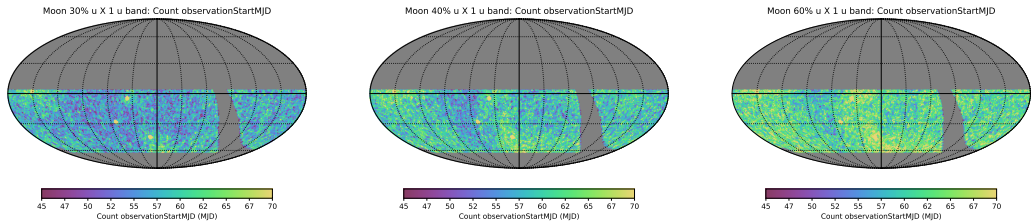

Figure 8. The number of visits in u band, with the filter load/unload at 30, 40 and 60% lunar illumination. This is with the u band survey footprint set to the standard weight; ideally the number of visits per pointing would be about 56. With a shorter period of time available for u band visits, the sky coverage is patchier. A filter swap at 40% lunar illumination does a reasonable job of achieving the required number of visits fairly uniformly across the sky.

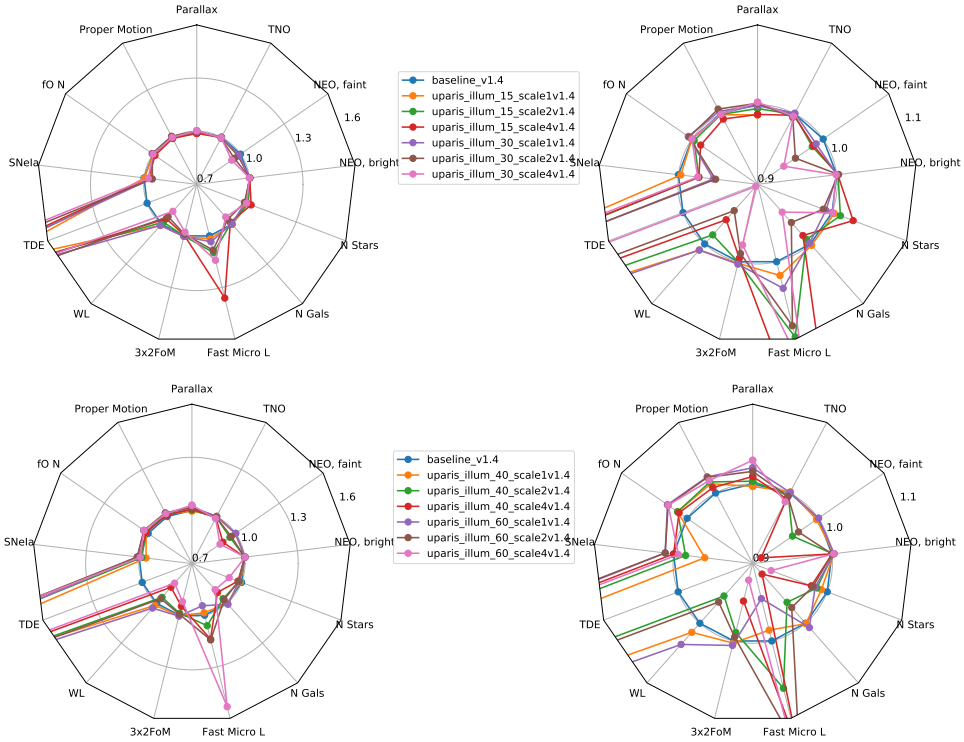

Figure 9. Varying the u filter load/unload time as well as the weight in the u band of the survey footprint. The right-most panels are zoomed in (around the baseline) versions of the panels on the left. 1

4.2. *FBS 1.5* `baseline`*: Baseline simulations (snaps and pairs)*

We use the baseline simulation as the touchpoint for the other simulations; the baseline serves as the reference for metrics and also sets a variety of default parameters carried into the other simulations.

This baseline survey is configured with 1x30s exposures per visit, with most visits ('blob visits') obtained in pairs separated by about 22 minutes (combinations of $u+g$, $u+r$, $g+r$, $r+i$, $i+z$, $z+y$, $y+y$ in any order for the pair). The footprint for the survey is the standard WFD plus minisurveys in the GP, NES and SCP, with five DD

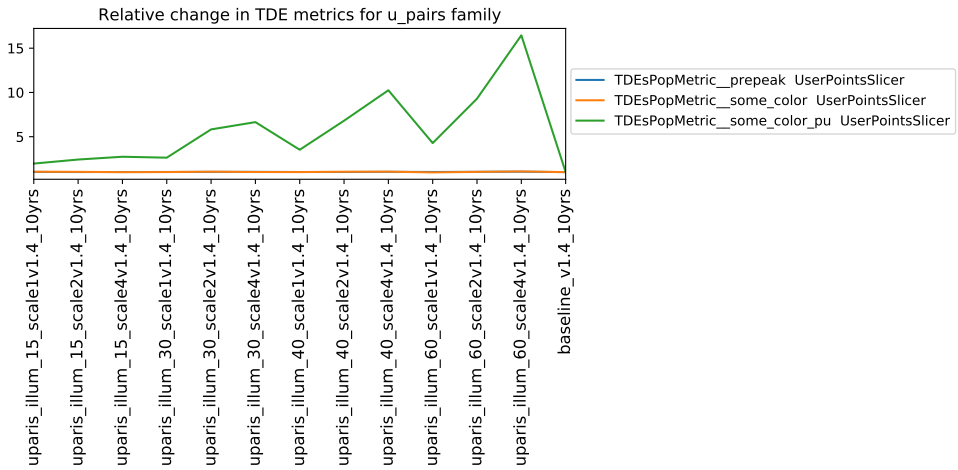

Figure 10. The relative change in TDE metric results across the family of *u_pairs* runs, compared to the baseline for FBS 1.4 (the FBS 1.4 baseline did not have u band paired with g or r; in `baseline_v1.4_10yrs` the u band was taken in single visits). There are multiple versions of this metric, corresponding to simple detection pre-peak, detection in any color, and detection in a color that includes u band; the final criteria is the most variable and the hardest to meet. Increased availability of the u band helps (up to 4x), and increasing the u band number of visits boosts the metric result as well (roughly linearly with the number of u band visits).

fields (located at the positions in Table 2). The ratios in the survey footprint between the various filters closely matches the desired distribution of visits over filters given in the SRD (u=6%, g=9%, r=22%, i=22%, z=20%, y=21% compared to an example SRD distribution of u=7%, g=10%, r=22%, i=22%, z=19%, y=19%). The u band filter was loaded in and out of the camera at 40% lunar illumination.

We also ran a comparison baseline simulation using 2x15s exposures per visit, instead of 1x30s. The overheads of taking 1x30s exposure per visit instead of 2x15s exposures per visit represent about a 9% decrease: 31 seconds per visit compared to 34 seconds per visit (with a single 1s shutter open/close, instead of 2s readout and 2x1s of shutter time, assuming the final readout occurs during the slew to the next field). This is reflected in the total number of visits acquired in each of these runs; there are about 8% more visits in `baseline_v1.5_10yrs` compared to `baseline_2snaps_v1.5_10yrs` It is worth noting that not all metrics scale directly with the number of visits; however, many do. Unfortunately, we *cannot assume* that 2x15s visits will not be necessary until the camera is on the telescope and the impact of cosmic rays and other artifacts are evaluated. Therefore, the correction between 2x15s visits and 1x30s visits should be kept in mind throughout the remainder of this work, even though all other simulations use 1x30s visits to evaluate the 'most likely' scenario.

The baseline simulation uses filters in mixed pairs; we did run a similar baseline-style simulation with 1x30s visits where the pairs were in the same filter (`baseline_samefilt_v1.5_10yrs`. This provides an efficiency boost due to fewer

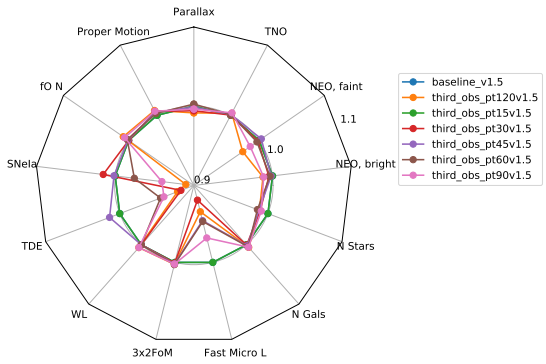

Figure 11. The science impact of dedicating the end of the night to gathering observations of areas that already have pairs. Note the scale - this radar plot covers 10% changes.

filter changes during the night, allowing on the order of 4% more visits over the lifetime of the survey. Observing in the same filter is beneficial for detecting solar system objects (since the limiting magnitudes of the pair of visits improves the likelihood of having detections for moving object linking), but is generally detrimental for measuring colors for transients and variables; for transients and variables, however, obtaining a third visit in the same night can be even more beneficial. A wider discussion of the intra-night cadence is covered in Section 6.2.

4.3. *FBS 1.5* `third_obs`*: Third Observation*

For early identification of transients, it can be helpful to have more than two observations in a night. Having two visits in the same filter, with a third in another filter, provides both a measurement of brightness variation over a short period of time and a color; this aids in classification. In this family of simulations, we dedicate between 15 minutes (simulation with the shortest time spent on triplets) and 120 minutes (simulation with the most time spent on triplets) at the end of each night to attempting to revisit areas of sky that already have been observed with a pair, in order to obtain a third visit. The greater the amount of time dedicated to this third observation, the less area of sky is covered in a given night. In general, the science impact of adding third observations seems to be fairly minimal or negative, see Figure 11. This is likely due to two effects: the metrics we're currently tracking aren't that sensitive to the presence of a third visit in a night (the TDE and SNIa metrics have negative impacts due to the lesser area covered per night, as they do not require three visits in a night), and highlights a need for a metric sensitive to this effect and appropriately tuned to highlight transient and variable classification and characterization requirements, and also triplets and quads exist in the baseline pairs simulations already, due to field overlaps. A wider discussion of the intra-night cadence is covered in Section 6.2.

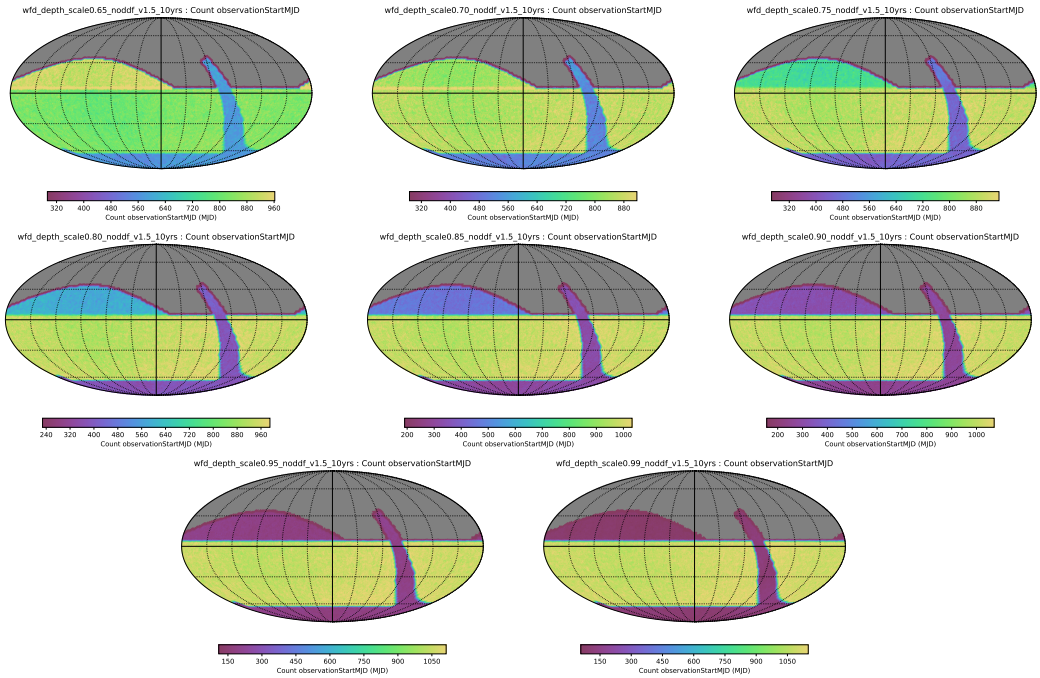

Figure 12. Varying the amount of time dedicated to the WFD region between 65% and 99% of the visits.

4.4. *FBS 1.5* `wfd_depth`*: WFD Weight*

This family of runs was primarily executed to confirm how the fO SRD metric scales with the footprint emphasis on the WFD. The survey footprint varies the fraction of observing time dedicated to the WFD area from 60% to 99%, with and without the standard DDF survey. For simplicity, here we look at the metric outputs run on the versions without the standard DDF fields; the numbers of visits per pointing that result are shown in Figure 12.

From these runs, we find that varying the fraction of time devoted to the WFD impacts various science metrics (see Figure 13), but even more importantly, it is likely that the SRD metric evaluating the minimum number of visits per pointing over the best 18k square degrees (fONv MinimumNvisits) cannot be met unless at least 70% of the survey time (in the footprint, which translates to more like 73% of actual visits due to dithering over the edges of the WFD region) is dedicated to the WFD. In these simulations, 73% of visits is approximately 1.65M visits out of the total 2.22M; in the case of bad weather, this would mean more visits would have to be redirected to the WFD.

In the remainder of our simulations, the amount of time dedicated to WFD varies, depending on the details of the survey footprint and minisurvey requirements. In general, it ranges from 66% to 94%, with most simulations falling around 83%.

4.5. *FBS 1.5* `footprint`*: WFD Footprints*

The location of the WFD region (and its filter distribution) are important questions for the LSST; the bulk of science from the LSST is expected to be facilitated by the

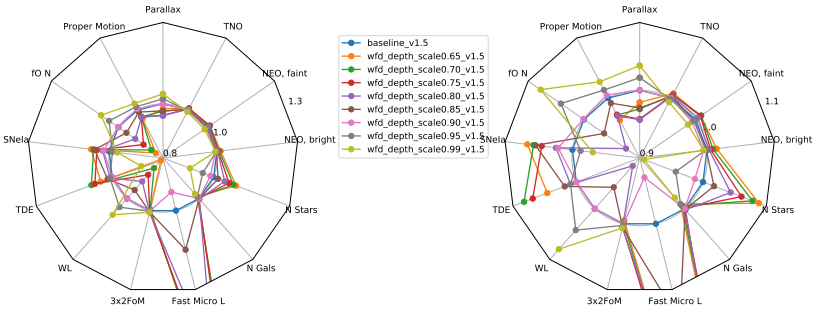

Figure 13. The science impact of varying the WFD depth. The right panel is a zoom-in on the left panel.

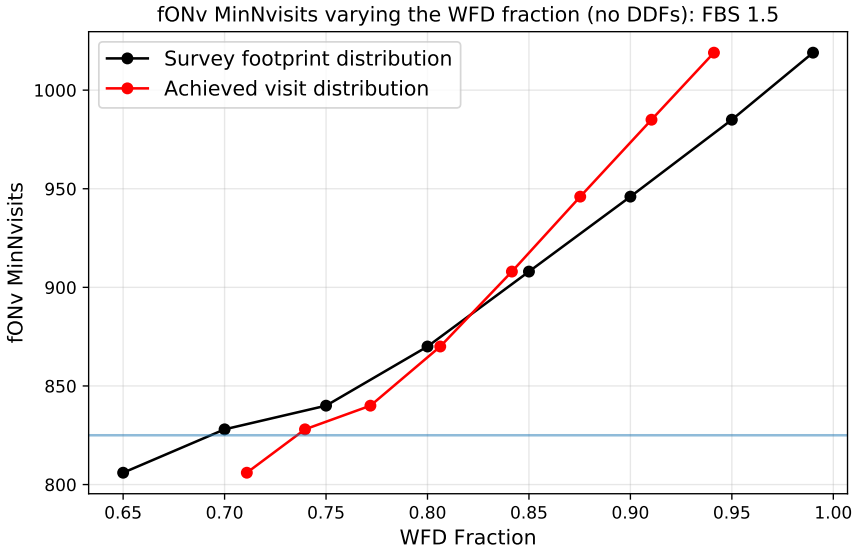

Figure 14. The minimum number of visits per pointing over the 'best' 18k square degrees (the WFD footprint), fONv MinVisits. When looking at the effect of scaling the WFD number of visits with a consistent footprint, this metric is perhaps more useful here than our standard version, the median number of visits per pointing over the WFD (fONv MedNvisits) as it will also capture 'patchiness' of visits. The black line shows the weight on the WFD in the survey footprint input to the scheduler; the red line shows the number of visits assigned to count toward the 'WFD area' out of the finished pointing history. The red line differs from the black line because entire visits were assigned as 'WFD', even when part of the field of view of the visit was beyond the 18k square degrees of the WFD. A 5% increase in time allocated to the WFD can result in every field receiving 40-50 more visits.

WFD. The WFD must be at least 18,000 square degrees to meet SRD requirements, however the location of those 18,000 square degrees is not specified. The standard baseline WFD includes regions which have dust extinction with E(B-V)$>$ 0.2. This amount of dust extinction is problematic for extragalactic science for two reasons – it reduces the effective coadded five sigma depth, and the total amount and wavelength dependence of dust extinction is not necessarily well characterized, so the effect on the background galaxies is hard to calibrate. An alternate 'big sky' WFD footprint

extending further north and south, but avoiding the galactic plane by a larger amount (either limited by dust extinction or by galactic latitude), can provide an 18,000 square degrees suitable for extragalactic science and moves parts of the NES and SCP into the WFD, but leaves larger amounts of sky to be covered toward the galactic plane in a separate mini-survey. See Figure 15 for more details. We ran several experiments with various survey footprints, some of which are more practical than others. The footprints in this section which leave no coverage of the galactic plane will be extremely detrimental to science which requires the galactic plane. Other footprints such as the newA footprint attempt to distribute coverage over too wide an area on the sky and so fail the requirements of 825 visits per pointing within the WFD. Many science metrics are extremely sensitive to the footprint; other families, such as the filter distribution family (Section 4.7) and the bulge coverage family (Section 4.6) are also important to consider as part of the overall footprint evaluation. A basic summary of the footprints in this section is shown in Figure 16; a wider discussion of the survey footprint is covered in Section 6.3.

4.6. *FBS 1.5* `bulges`*: Galactic plane coverage*

The survey strategy for the galactic plane is an important question for science dealing with populations within the Milky Way, especially transients and variables that are most populous in the plane and toward the Magellanic Clouds. The SAC made a series of recommendations for survey strategy in the galactic plane, which were implemented in this family of simulations. The background WFD footprint for this family is the 'big sky' style footprint introduced in the previous section.

We use three footprints for bulge coverage:

- light coverage of the bulge and entire galactic plane (`bulges_bs`),
- the galactic bulge as deep as WFD (`bulges_bulge_wfd`),
- the galactic bulge covered similarly to WFD, but with more observations in i (`bulges_i_heavy`.

See Figure 17 for more details on the location of the increased plane coverage. For each of these strategies, we run a version with natural cadence and one where we boost the priority of the bulge if it has not been observed in 2.5 days (to ensure a more rapid cadence).

Covering the bulge understandably increases the overall number of stars expected from the simulation, as well as the fast microlensing events (which are primarily concentrated toward the bulge). Because this requires more visits away from the larger 'big sky' region of WFD, there is a slight decrease in the SRD metrics. See Figure 18.

4.7. *FBS 1.5* `filter_dist`*: Filter Distribution*

The filter distribution for the standard baseline simulation follows the suggested distribution in the SRD, however this family varies the weights between different filters

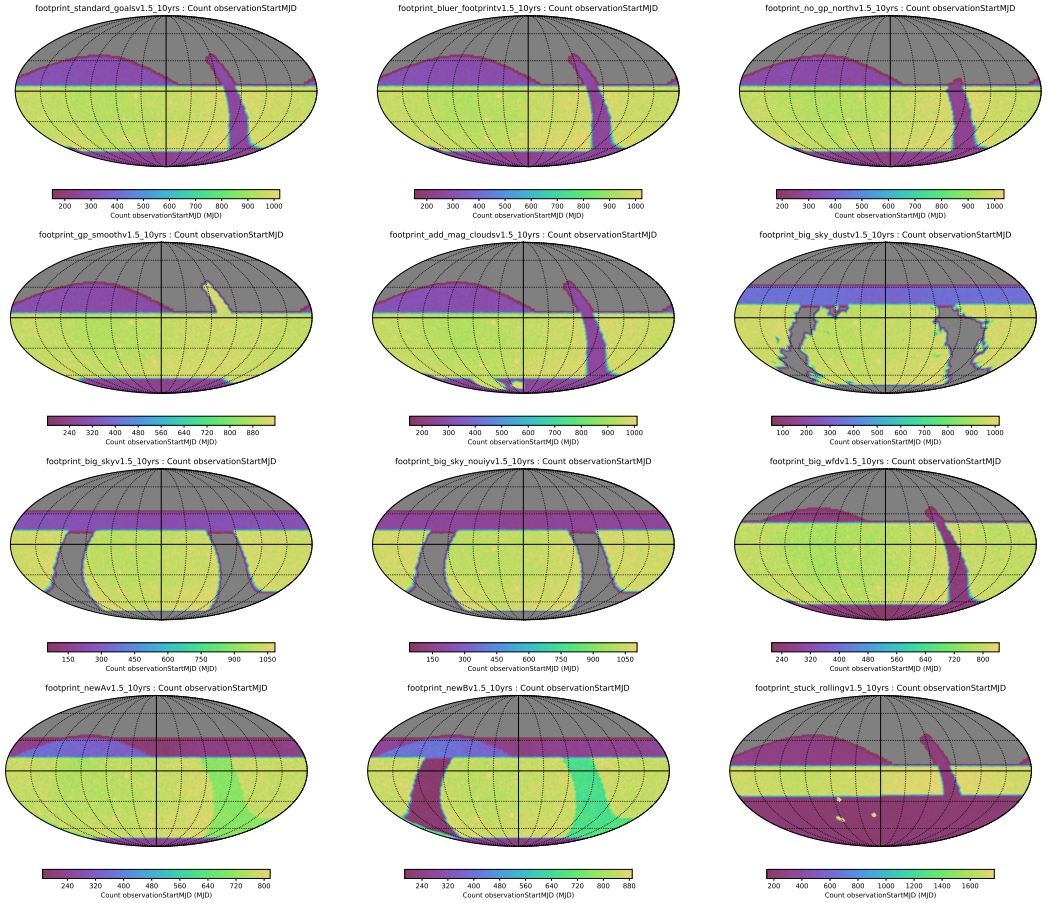

Figure 15. The total number of visits in each of the different survey footprints simulated. From top left to right and then down, the description of these survey footprints is:
(1) standard (previous) baseline survey footprint (`footprint_standard_goals`),
(2) the same footprint but with a bluer filter distribution (`footprint_bluer_footprint`),
(3) the standard footprint but removing the northern tip of the galactic plane minisurvey (`footprint_no_gp_north`),
(4) the standard footprint, but continuing the WFD-cadence of visits through the galactic plane (`footprint_gp_smooth`),
(5) the standard footprint but adding a magellanic clouds extension at WFD-cadence (`footprint_add_mag_clouds`),
(6) an extended N/S footprint (going about 10 degrees further north and south) nicknamed 'big sky', with the galactic plane boundaries delineated by dust extinction, a small northern extension but no SCP or GP coverage `footprint_big_sky_dust`),
(7) an extended N/S footprint similar to the previous, but with the galactic plane boundaries delineated by galactic latitude ($l = 20$) (`footprint_big_sky`),
(8) the same footprint, but without any coverage in u, i or y band (`footprint_big_sky_nouiy`),
(9) an extended WFD region, going further north in the sky (even further than big sky) although not as far south, includes SCP and GP coverage with a small extension for the NES (`footprint_bigwfd`)
(10) an extended N/S WFD footprint (in the 'big sky' style) but with the galactic plane defined by galactic latitude ($l = 20$) – this adds the GP covered to just slightly less than WFD depth, and minisurveys for the SCP and NES (`footprint_newA`)
(11) similar to the newA footprint, however the galactic anti-center is covered with fewer visits, to allow more visits over the WFD region (`footprint_newB`)
(12) this survey footprint is primarily a test case to find wide area metrics that were not properly sensitive to area; here the WFD is purposefully not covered appropriately but rather the northern half of the standard WFD received almost all of the visits from the WFD while the southern half receives a small fraction (`footprint_stuck_rolling`).

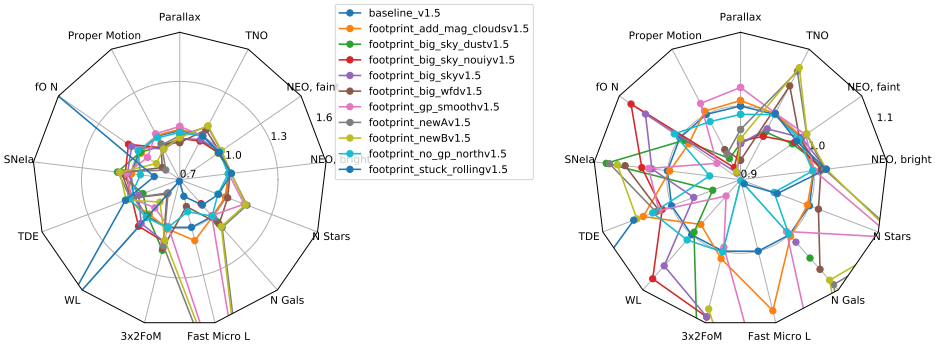

Figure 16. Science impact of varying the WFD survey footprint. The number of stars and galaxies is obviously very sensitive to the footprint, as are the number of discovered TNOs (as these objects move very slowly). The Fast Microlensing also varies strongly, as this metric depends on galactic plane coverage. The right panel is a zoom-in of the left panel.

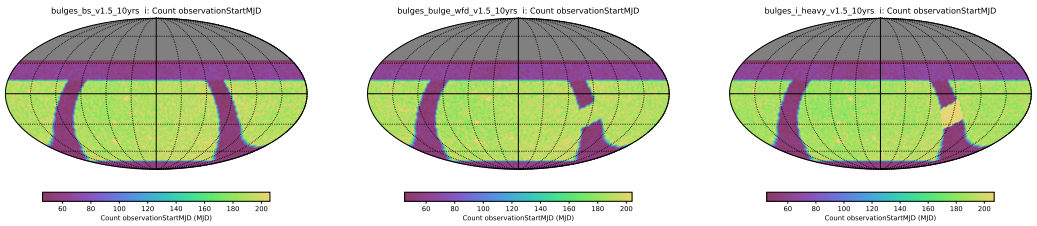

Figure 17. Series of simulations trying different bulge observing strategies; these vary from simple light coverage of the galactic bulge and plane to heavier coverage of the bulge (while maintaining light coverage of the rest of the galactic plane).

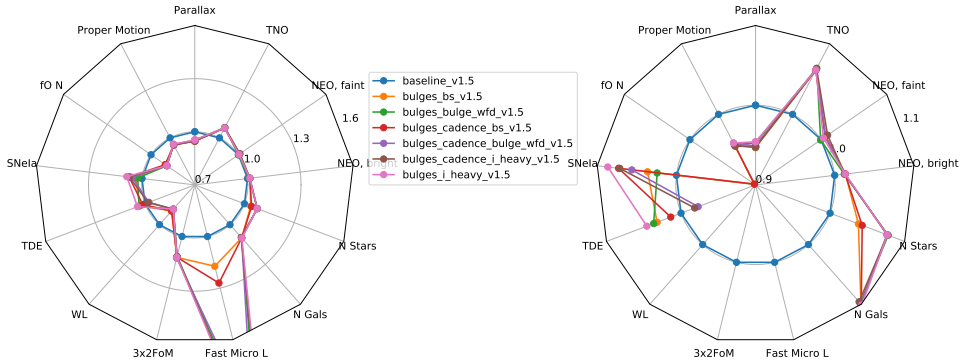

Figure 18. Science impact of the different bulge strategy simulations. The right panel is a zoom in of the left.

significantly. This family should serve as a useful testbed for photometric redshift evaluations, but we do not currently have a photometric redshift metric available. However, this family does also illustrate other tensions between SNe and solar system science, for example, with SNe benefiting from more visits in bluer filters while solar system discovery (because these objects are red) prefer more visits in redder bands. This family is also useful when evaluating the overall survey footprint, as it is a simple WFD-only survey footprint, with no SCP or (significantly for solar system objects)

Name	u	g	r	i	z	y
Uniform	1.00	1.00	1	1.00	1.00	1.00
Baseline	0.31	0.44	1	1.00	0.90	0.90
g heavy	0.31	1.00	1	1.00	0.90	0.90
u heavy	0.90	0.44	1	1.00	0.90	0.90
z and y heavy	0.31	0.44	1	1.00	1.50	1.50
i heavy	0.31	0.44	1	1.50	0.90	0.90
Bluer	0.50	0.60	1	1.00	0.90	0.90
Redder	0.31	0.44	1	1.10	1.10	1.10

Table 3. Variations of the filter distribution simulated.

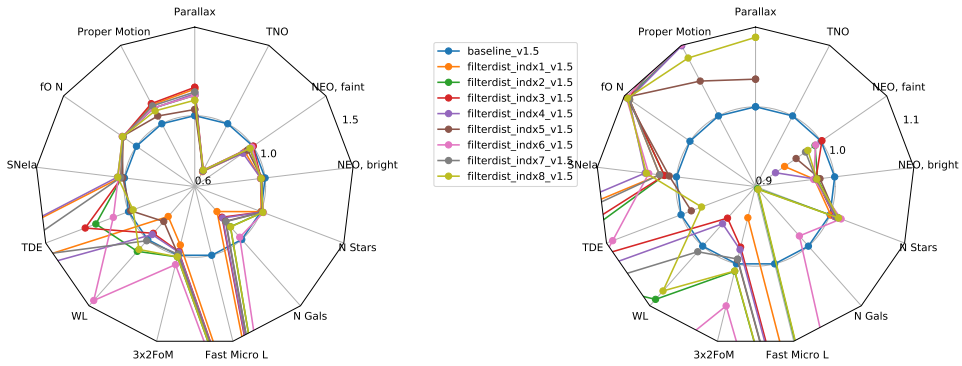

Figure 19. Science impact of varying the filter distribution

NES. The filter distributions simulated are listed in Table 3, and the radar plot of high level metrics is Figure 19.

4.8. *FBS 1.5* `alt_roll_dust`*: Nighty N/S alternating observing*

This family of simulations pulls in a feature from the altSched simulations (D. Rothchild et al. 2019), where visits alternate between northern fields and southern fields on a nightly basis. It uses the big sky dust-extinction limited footprint, and adds a basis function to encourage the scheduler to alternate between the north and south nightly. This can help keep light curve sampling optimally spaced, but does induce (at least) a night-long gap between revisits to a field (See Figure 21). As this basis function is not as strict as the scheduling in the altSched simulations, the telescope avoids pointing too close to the moon. One of these simulations simply alternates N/S on alternate nights; the other adds a modified 2-dec-band rolling cadence, where the alternating N/S visits are maintained within the rolling declination bands.

There is no additional NES, however there is a strip in the north observed in g, r, i, and z. The survey footprint focuses on low dust extinction regions and includes the galactic bulge; it covers more area than the standard baseline, so observes more stars and galaxies. The coverage of the LMC and bulge increases the number of fast microlensing events.

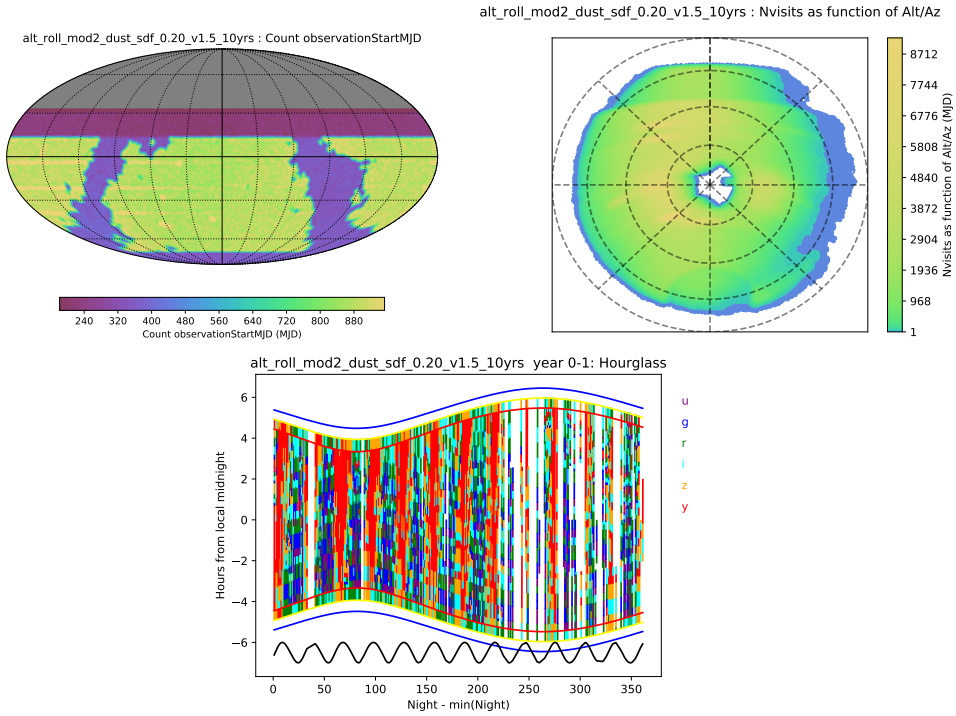

Figure 20. The `alt_roll_dust` simulation uses a footprint to avoid high dust extinction and tries to drive an every-other-day cadence.

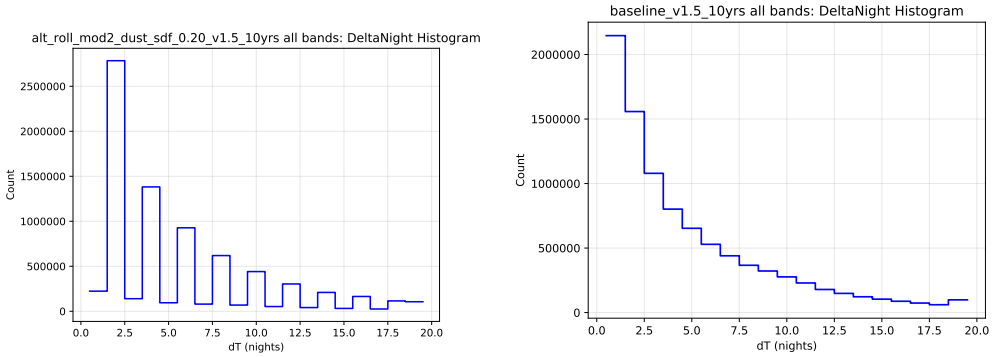

Figure 21. Histograms of the inter-night revisit intervals for the `alt_roll_mod2_dust` simulation (left), with alternating N/S visits on alternating nights, compared to the same histogram for the `baseline_v1.5` simulation (right).

The science impact of this strategy is fairly minimal. By avoiding extinction regions, we have more stars and galaxies. The coverage of the LMC also increases the number of fast microlensing events. The SRD metrics are are lower than baseline, due to a lower fraction of visits being focused in the WFD, but these still meet requirements. See Figure 22.

4.9. *FBS 1.6* `rolling_fpo`*: Rolling Cadences*

Rolling cadence is the term we have given to executing the survey in a non-uniform manner, emphasizing some region of sky one year, then deemphasizing it the next.

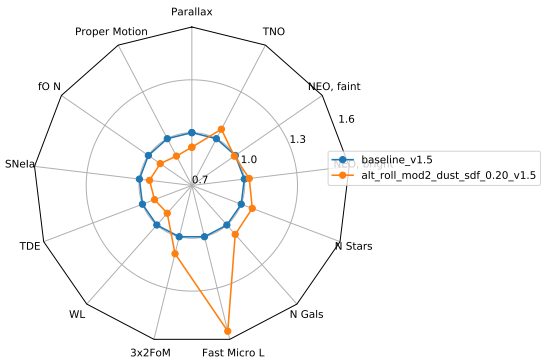

Figure 22. The science impact for alt_roll_dust.

Because the SRD includes requirements on stellar proper motion measurements, we are constrained to cover the sky uniformly in at least year 1 and year 10. We experiment with using rolling cadences where the WFD region is divided in 2, 3, and 6 declination bands (see Figure 23). We also scale the rolling weight to be 80, 90, and 99%; a larger weight results in more visits in the emphasized declination band and fewer outside the band. It may be reasonable to expect some visits over the entire sky in each year, if templates for image subtraction require these visits (and if ToO programs require updated templates), and there may be additional benefits for other science requiring long-term photometric monitoring.

These simulations were created using the FBS 1.6 code; this version of the scheduler uses an improved method for determining the length of time a field may have been available for observation when calculating the reward for the footprint map. The result of this upgrade is smoother, more even rolling cadences. The survey footprint, aside from rolling, was the standard baseline footprint.

Figure 24 shows the science impact of the different rolling cadence simulations. In general, the rolling cadence is fairly neutral for these metrics, at the 2 or 3 declination band level. The 6-declination-band rolling cadence can have negative effects on several of the science metrics, which seems likely to be linked to the limited amount of area covered within each band – the increased cadence is more frequent than required, but the amount of sky covered has dropped, so most metrics sensitive to both area and cadence will have a negative response.

4.10. *FBS 1.5* `ddf`*: Deep Drilling Fields*

The locations for four of the Deep Drilling Fields has been determined for some time; the location of the fifth field is now planned to overlap with the Euclid South field (two LSST pointings are required to actually match the full Euclid South field; we are currently simply alternating between two field centers in this location). The locations of the DDFs used in these simulations are listed in Table 2. and can be easily seen in the overall survey footprint map in Figure 25. The cadence of visits,

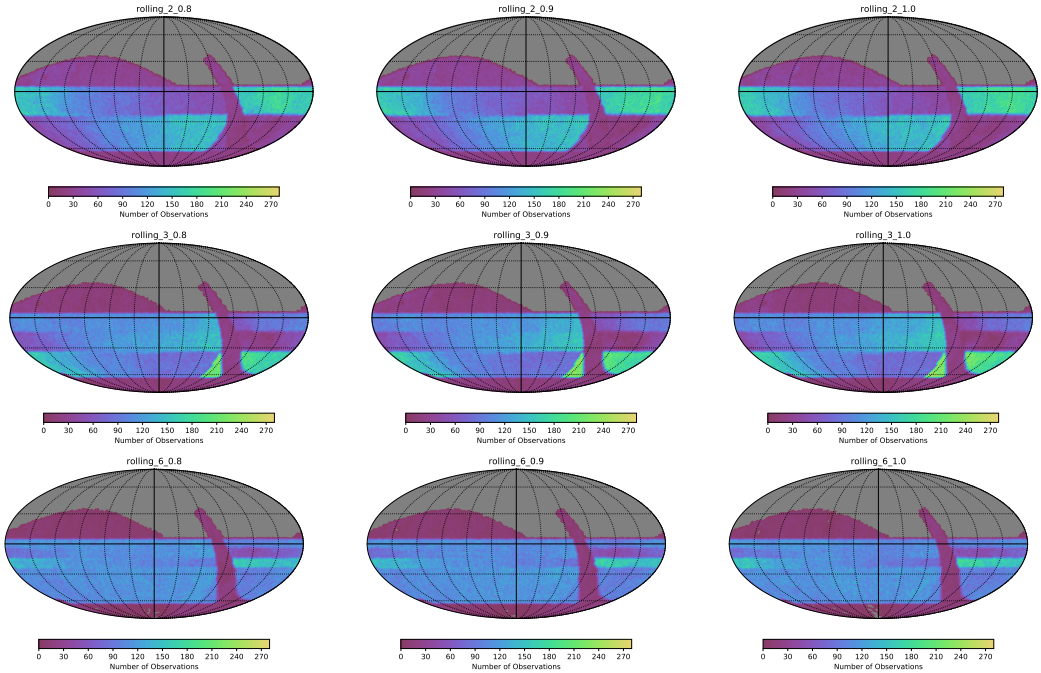

Figure 23. Rolling cadence simulations with 2 (top), 3 (middle), and 6 (bottom) rolling stripes. Here we show the observations taken from 3.5-4.5 years in the survey, excluding the DDF observations.

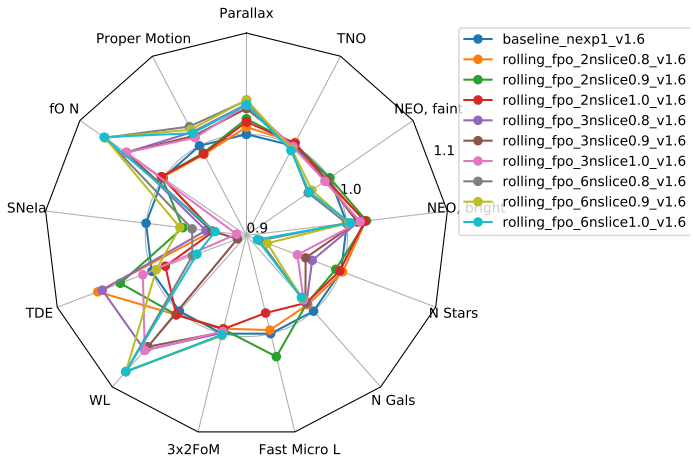

Figure 24. Science impact of different rolling simulations. The 6-band rolling cadence has some negative science impacts, while the 2 or 3 band cases tend to be fairly neutral. It seems likely that more metrics focusing on transients and variables will be useful here.

as well as coadded depths of these DDFs, still needs to be finalized; the cadence has implications for the overall amount of time required for the DDFs.

The DDF families of simulations test different cadences for the DDFs, with a standard baseline survey footprint and strategy for the rest of the sky. The DDFs are somewhat decoupled from much of the remainder of the survey strategy; once given a fixed fraction of total observing time, the scheduler will keep DDF visits within the allocated fraction at all times, and the start or end of DDF visits don't interfere with

other observations (the scheduler waits until a blob is finished before starting a DDF sequence).

Within these bounds, we ran a variety of DDF strategies, based on requested cadences in the white papers and some test extensions:

- AGN: This strategy takes shorter DDF sequences more often. Only ∼2.5% of visits are spent on DDFs, making the final coadded depths much shallower than other strategies.
- DESC: a strategy that split the blue and red filters to different days, emphasizing a 3-day cadence. The overall time request is about 5%.
- Baseline: Our baseline strategy where 5% of observations are allocated to DDF observations. Sequences include whichever filters are available out of *ugrizy* and then take *u*x8, *g*x20, *r*x10, *i*x20, *z*x26, and/or *y*x20 visits per band, all with 30s exposures
- Daily: Similar to the baseline, but includes shorter DDF sequences that can execute daily so there are no long gaps between observations. Total DDF fraction is 5%.

Figure 27 shows the same observing season of the DDF ELIASS1 with these different strategies.

Figure 26 shows the different coadded depths (left) and science impact (right) of the variations on the DDF strategies we have tried. If the DDFs are kept to a consistent fraction of time, the overall science impact tends to be consistent.

The fraction of time required for the DDFs is driven by the cadences required – the season length, the inter-night cadence (number of nights between sequences), and the number of visits and filters required in each sequence. The sequences for the AGN strategy and the sequences for the DESC strategy are not completely compatible; an attempt to satisfy both cadence requirements will need more than 5% of the total time, and perhaps as much as 8% of the total time. We are missing metrics to evaluate the science performance of each of these DDF strategies, so it is unclear how well these simulations meet the desired goals.

4.11. *FBS 1.5* `good_seeing`*: Good Seeing Images*

In order to obtain good seeing images for difference imaging templates and to improve a host of other image processing issues such as deblending, it may be desirable to ensure that the entire WFD area is imaged in 'good seeing' conditions every year. We defined 'good seeing' as FWHM of 0.7 arcseconds or better, then set up simulations where we required one good seeing image in various bandpasses each year.

The science impact of adding this relatively simple constraint to the observing strategy seem minimal and generally positive (see Figure 28). We do not currently have a metric tied directly to seeing, although of course image depth is important to many metrics and this depends on seeing.

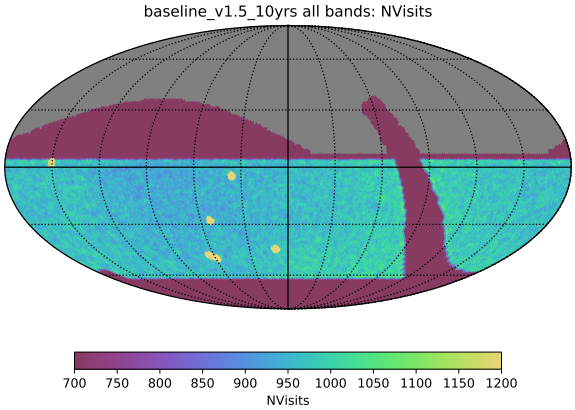

Figure 25. Number of visits per pointing on the sky, in all filters, for `baseline_v1.5_10yrs`; the locations of the five DDFs (with the double-pointing for Euclid South) are easy to pick out. The locations of these DDFs remains the same in all simulations.

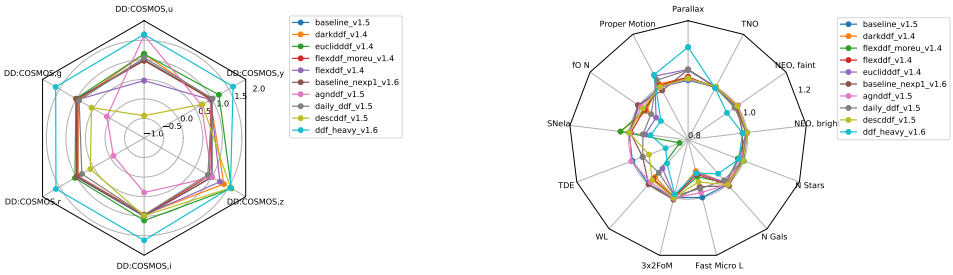

Figure 26. On the left, we show the coadded depth in each filter for a representative Deep Drilling Field. Larger values mean deeper coadded depth (units are magnitudes). On the right we show the standard science metrics, evaluated per DDF.

There is no obvious additional overhead to observing, although this may be slightly more challenging to implement in operations in terms of generating and passing a queue from the FBS to the telescope. While the FBS can simulate an entire night and then pass this to the observing queue, if seeing conditions are highly variable the queue may need to be regenerated more frequently. This is likely an issue to address with additional telemetry from the site.

4.12. *FBS 1.5* `twilight_neo`*: Twilight NEO Survey*

At the time of the call for white papers, we were using a more optimistic weather cutoff and so had more time available for observing; using twilight time for non-WFD purposes seemed an ideal option. With the more conservative weather cut we are currently using, and the number of minisurveys and DDFs running concurrent with the WFD, we need to use at least some of the twilight time for WFD visits (and indeed, have been in previous simulations). Twilight observing starts at -12 degree solar altitude, continuing to about -19 degrees; nautical to astronomical twilight.

The Seaman et al. white paper suggested a twilight survey for PHAs and NEOs, looking at low elevations (high airmass) near the sun. These can be highly productive

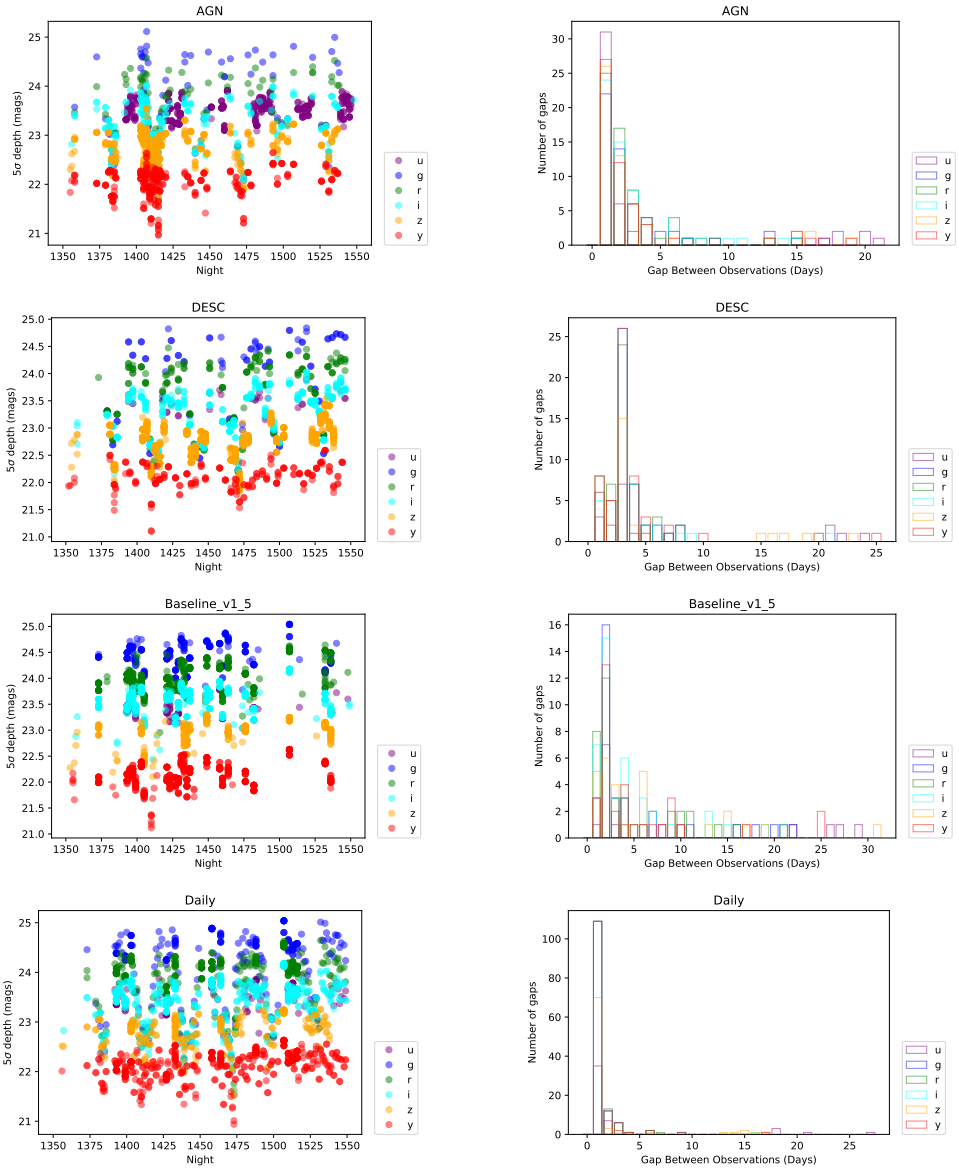

Figure 27. One observing season of the DDF ELIASS1 under different DDF strategies. See Section 4.10 for more information.

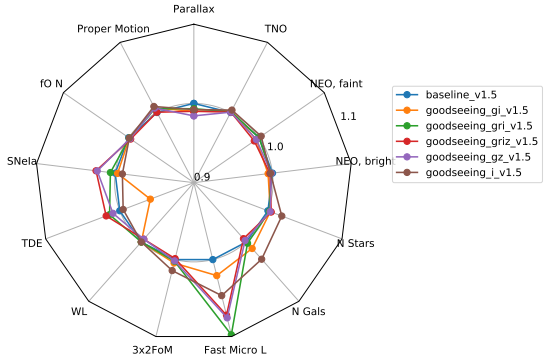

Figure 28. The science impact of making sure the sky has template images in good seeing conditions.

options for many surveys, covering the NEO 'sweet spots' and are the only kinds of visits that have the capability to detect Vatiras, asteroids with orbits interior to Venus (Note: the solar system metrics do NOT include a population of Vatiras currently).

The twilight survey, as implemented, adds a set of visits consisting of 1 second exposures in r, i or z bands (depending on the filter previously in the camera, so tied loosely to lunar phase). Observations are attempted in morning and evening twilight on the nights where the twilight survey is active, typically about 440 1s visits per night. Visits are acquired in triplets, to identify fast moving objects within a single night. The fields chosen for the NEO survey are within 40 degrees of the ecliptic, at high airmasses toward the Sun. In this family of simulations, the twilight survey is activated for varying fractions of time: either every night, every other night, every third night or every fourth night (roughly .. weather or downtime can interfere). The final sky coverage for the twilight_neo survey, if activated every available night, is shown in Figure 29.

We find that running the twilight survey every available night causes the main survey to fail to meet SRD requirements in the baseline survey configuration. The twilight time is needed for WFD and other minisurveys; with an every-night twilight neo survey, there are $\sim$ 10% fewer long exposure visits in the rest of the survey; with an every-other night twilight survey there are $\sim$ 5% fewer long exposure visits over the rest of the sky. Running the twilight neo survey less often has less of a negative impact, but it also has less of a positive impact on the completeness of the large PHAs and NEOs. The smaller PHAs and NEOs have a lower overall population completeness when the twilight neo survey uses more time; these objects are likely too faint to be discovered in the short exposures of the twilight survey, thus need to be discovered in the standard survey. Discovery for fainter solar system objects is dependent on the number of visits available (for a given footprint), so it is consistent that the final completeness for the smaller objects drops. See Figure 31.

An obvious missing metric for this survey is sensitivity to Vatira discovery (which means a new input population more than than a new metric, per se); we can check very approximately the sensitivity of this implementation of twilight observing for objects close to the Sun by looking at the distribution of solar elongation angles for the twilight visits (see Figure 32). The distribution of solar elongations is not as strongly concentrated toward the Sun as would be ideal for this purpose. It would be useful to extend this simulation family with twilight moving object surveys using slightly different strategies; instead of a fixed cadence throughout the month, the cadence should likely pause depending on the location of the moon (the high solar elongation visits in the current simulation are likely avoiding the moon), and slightly longer visits may be useful.

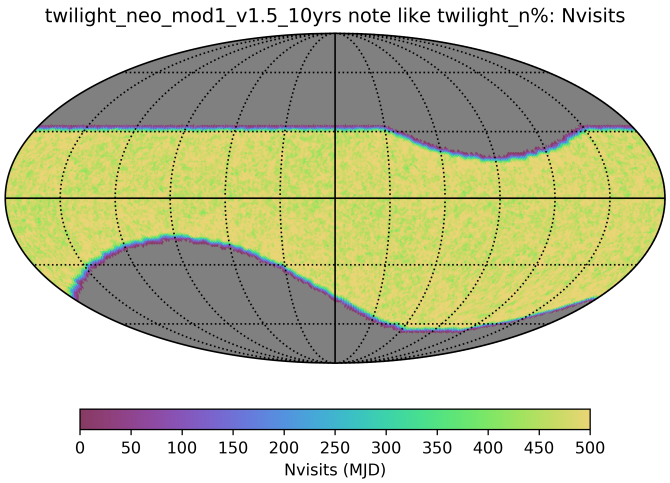

Figure 29. Sky coverage of the NEO twilight survey, at the end of 10 years. These visits are in r, i and z band, with 1 second duration.

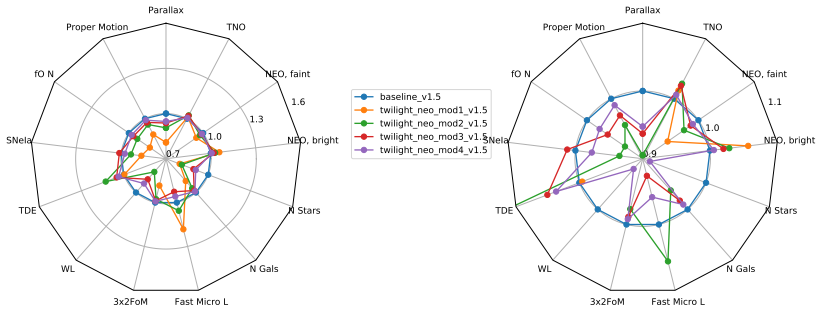

Figure 30. The science impact of using some or all of twilight time for a NEO survey.

4.13. *FBS 1.5* `short_exp`*: Short Exposures*

In this family of simulations, we added a minisurvey to obtain short exposures (either 1s or 5s) twice per year or five time per year, in all filters. Short exposures allow enable direct comparisons for photometry and astrometry to brighter catalogs and enable some science with bright objects which saturate in the standard visits. Taking shorter exposures is a less efficient observing mode, but it has little impact on the overall open shutter fraction; this makes sense given that the overall amount of time used for these short-exposure minisurveys is relatively small. The short exposure surveys do have a small but noticeable impact on the rest of the survey, just due to their time requirement - the total number of long (visit exposure time $>$ 15 seconds) visits drops in relation to the amount of time taken by the minisurvey, up to 7% for the 5s visits 5 times per year simulation (see Figure 34). This is reflected in small drops in most science metrics, Figure 35. The TDE metric, which can be very sensitive to small changes in u band visits and cadence, shows a larger drop; some of this is related to statistical noise in the metric.

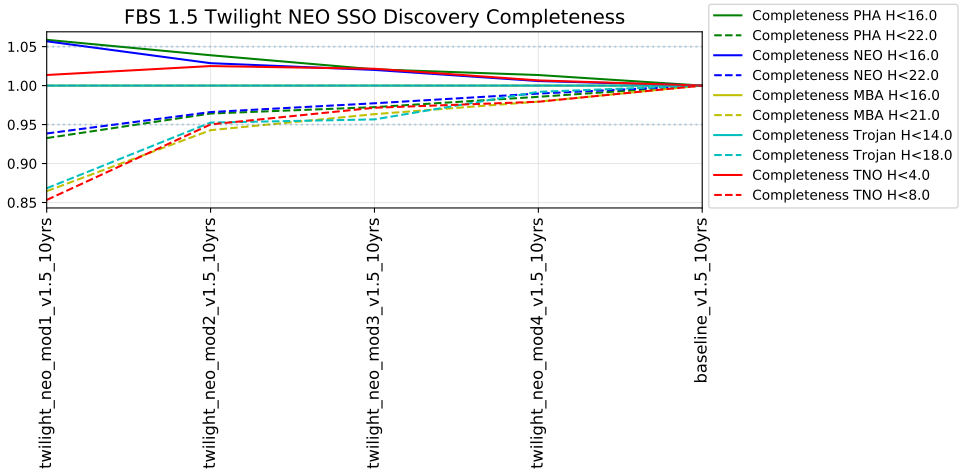

Figure 31. More detailed solar system metric results for large (solid lines) and smaller (dashed lines) members of the moving object populations. Large PHAs and NEOs benefit from twilight neo survey, and more time spent on the twilight survey boosts their completeness higher. Smaller objects, which are also fainter, have lowered completeness. This is likely because the smaller objects are not visible in the short, 1s exposures, yet the larger number of those short exposures lowers the total number of visits in the standard survey. The smaller objects are always more sensitive to the total number of visits, due to having fewer opportunities for discovery while they happen to be bright enough to make it beyond the limiting magnitude cuts.

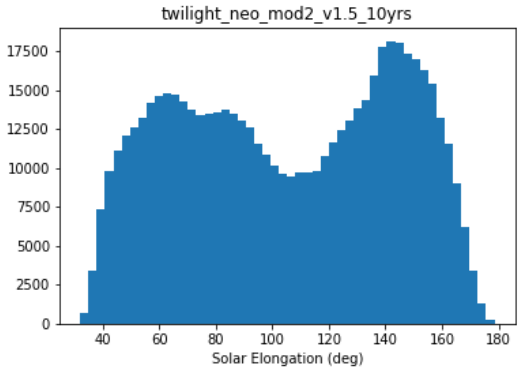

Figure 32. The distribution of solar elongations for the twilight NEO survey fields in `twilight_neo_mod2` simulation; the distributions are similar for the other versions of the twilight NEO simulations. The maximum solar elongation of Venus is 45 degrees; we could potentially detect objects interior to Venus's orbit with these visits.

In order to properly evaluate the benefit of these short exposure surveys, we need metrics which would be sensitive to the science enabled by bright objects measurable in the 1s or 5s visits.

4.14. *FBS 1.5* `u60`*: Longer u Exposure Time*

Due to the low sky background levels and measured levels of read-noise in the amplifiers, u band observations within the survey are often expected to be read-noise limited for either 15s or 30s exposures. Doubling the u band exposure time to 60

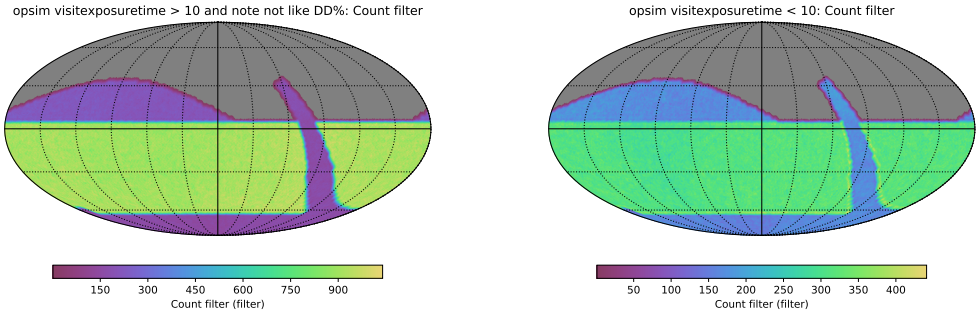

Figure 33. Results from including 5s exposures (up to 5 per year). The left shows the number of regular 30s visits (excluding DDF observations) and the right shows the number of 5s visits.

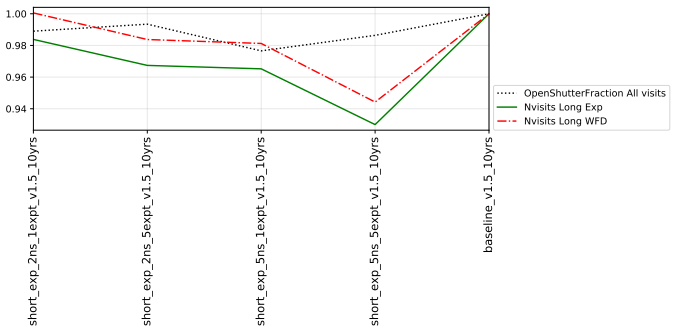

Figure 34. As the number of short exposure visits increases and the exposure time of the short exposure visits increases, the total number of long exposure visits in the rest of the survey drops by about 1 to 7%.

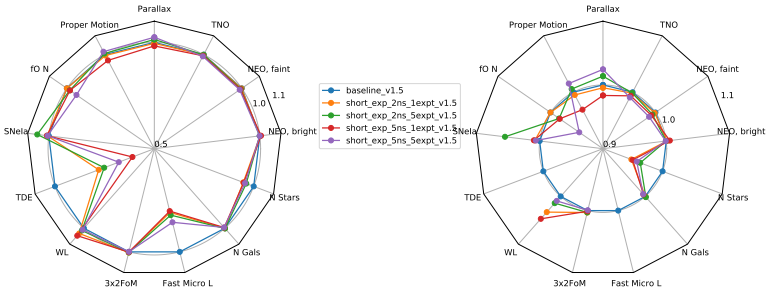

Figure 35. Science impact of covering the sky in short exposures.

seconds moves these back into the sky-noise limited regime. This simulation sets the u band visit exposure time to 60s (1x60s visits) and cuts the total number of u band visits in half. This increases the final u band final coadded depth by ∼0.20 magnitudes. An interesting side effect is that the g band coadded depth also increases by about 0.10 magnitudes; the g depth increases because 60s u visits decreases the overhead time (and consolidates u band into a shorter period of real time), which frees up more dark time for g visits. Other filters are essentially unchanged in final coadded depth. The total number of visits in the simulation drops by 6%; this could have been closer to 3% (halving the u band visits), but visits in g and r band which

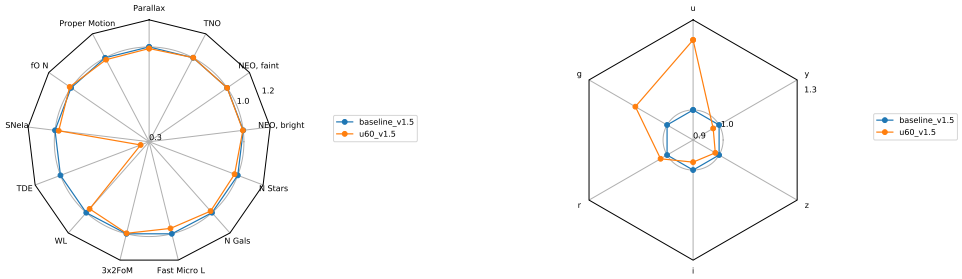

Figure 36. Science impact of ncreasing the u exposure time to 60s (left). As expected, this results in a substantial gain in u coadded depth (right).

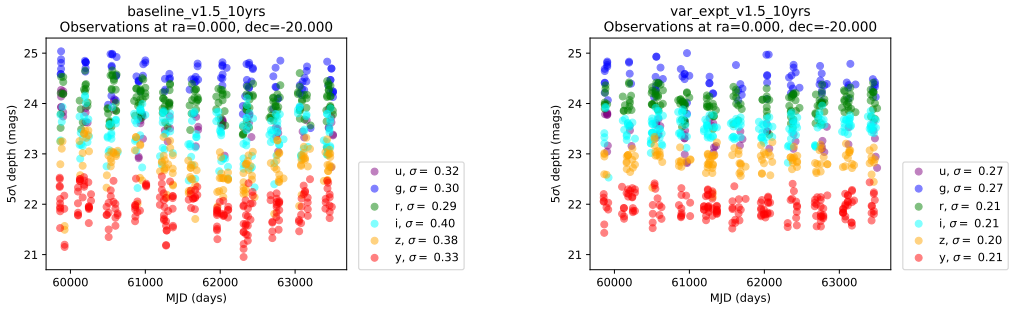

Figure 37. Comparison of a five sigma depths (for point sources) at an example location in the WFD, between the baseline (left) and the variable exposure time simulation (right). With variable exposure times, the individual observations five sigma limiting magnitudes become more uniform, especially in the redder filters that can be observed in bright time and twilight.

were paired with u band also were extended to 60s to maintain the time spacing between pairs of visits during each blob.

Consolidating u band exposure time into fewer visits is likely to be problematic for classification of many transient objects, as they will be less likely to have a u band visit which is often a feature used for distinguishing various kinds of objects. Note, we assume that 1x60s visit may count as 2 30s visits for the purpose of meeting the SRD value of 825 visits in the WFD area (the `u60` run will meet this requirement without this assumption, but depending on other survey requirements this may be an important assumption to adopt). See Figure 36.

4.15. *FBS 1.5* `var_expt`*: Variable Exposure Times*

In order to maintain more uniform limiting magnitudes in individual visits, it may be beneficial to vary the visit exposure times. This would mean that transients and variables detected in images during good conditions (dark sky, good seeing, near zenith, etc.) would be more likely to also be able to be detected in images taken in poorer conditions. It would also mean that during poor conditions, longer exposures would mean more useful information than if the consistent 30s visit time was maintained.

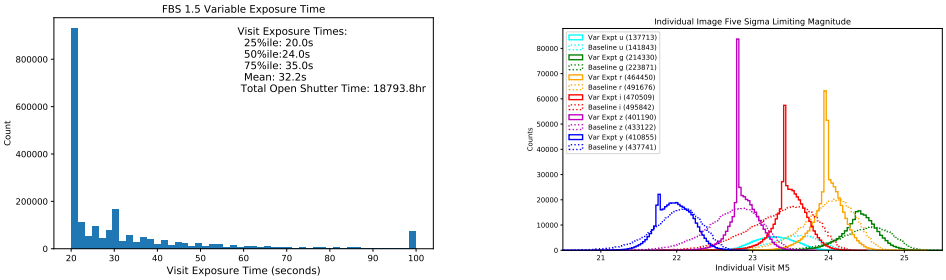

Figure 38. Distribution of visit exposure times in the variable exposure time simulation (left) and comparison of the individual image five sigma limiting magnitudes, including the baseline (right). In the baseline, there would be simply 30 second exposures; in the variable exposure simulation the mean visit exposure time is 32.2 seconds, but visits range from 20s to 100s. The variable exposure times cut down on the widths of the five sigma depth distribution, and tends to slightly increase the mean image depth (in most but not all bands). The variable exposure simulation contains 6% fewer visits than the baseline survey.

This simulation tests the effect of variable exposure times. In good conditions, the exposure time is allowed to shrink to 20 seconds, while in poor conditions it can extend to 100 seconds. This maintains an ideal depth which is somewhat shallower than the typical image depth in the baseline simulation, but is more consistent across visits. Figure 38 shows the distribution of visit exposure times and individual visit limiting magnitudes achieved in the simulation. The median visit time was 24 seconds, shorter than the baseline survey exposure times, but the mean visit time was 32.2 seconds – there are many visits with 100s exposure times, which drives the survey to 6% fewer overall number of visits. The scatter in the image depths was reduced by a tenth of a magnitude or so in most bands, although with a variable change in the median and mean five sigma depths (u became shallower while i and z gained some depth). The total on-sky exposure time is very similar to that of the baseline survey, but the coadded depths in u and g bands become shallower by 0.3 and 0.1 magnitudes, respectively (see the right panel of Figure 39).

Having variable exposure time introduces at least 8 new free parameters to the scheduler: the target individual depth for each filter, as well as the shortest and longest acceptable exposure times. As with 4.11, this would be more complicated to run in operations as the scheduler would need current conditions to calculate the modified exposure times, although the predicted sky brightness may be accurate enough.

As with doing 60s u band exposures, having longer visits will reduce the overall number of visits; longer visits may need to be allowed to count as multiple visits for the purpose of meeting the SRD requirement of a median of 825 visits per pointing over 18k square degrees (fONv MedVisits). Figure 39 shows the science impact of varying the exposure time is fairly minimal.

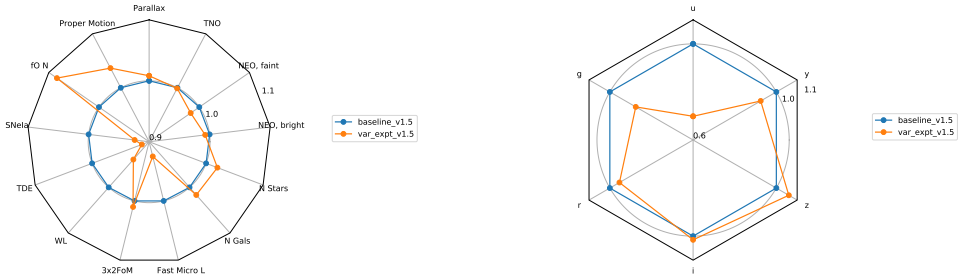

Figure 39. Science impact of using variable exposure times (left). The final coadded depths also shift, with decreases in final depth in the u and g bands and increases in z (right).

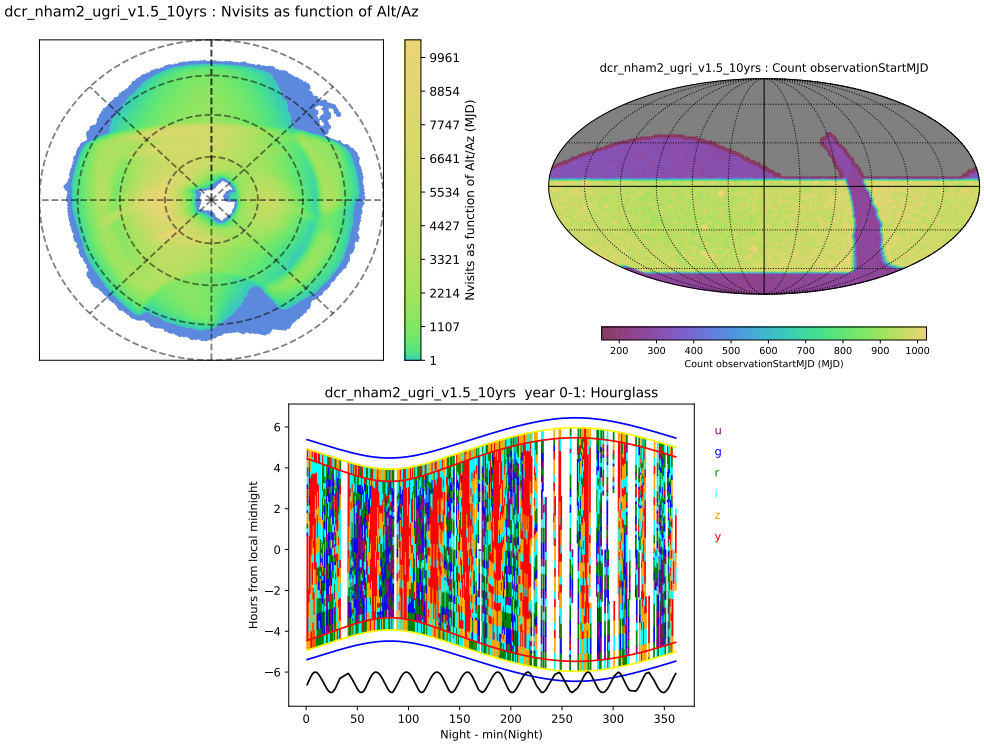

Figure 40. Intentionally taking observations at higher airmass to measure DCR.

4.16. *FBS 1.5* `dcr`*: DCR visits*

A consistent problem in difference imaging are dipoles resulting from differential chromatic refraction (DCR); this is potentially a larger problem for the LSST as the telescope does not have an atmospheric dispersion corrector and will have occasion to take images at high airmasses. However, if images are obtained to measure the effects of DCR, then this effect can be corrected for when building and applying the template images used in difference imaging. It is also possible to use the measured chromatic shift in objects with sharp features in their SEDs (e.g., quasars with strong emission lines) to measure properties of those SEDs, providing potential science opportunities.

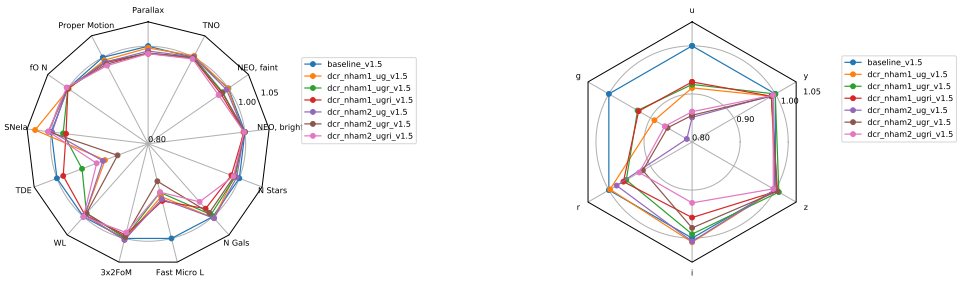

Figure 41. Science impact of including observations at high airmass for DCR. As expected, pushing observations to high airmass lowers the coadded depths (right) and has as slight negative impact on most science metrics (left).

This family of simulations intentionally schedules a subset of images at high airmass to enable building a DCR model. We test adding various combinations of filters to these high airmass, DCR visits ($u+g$, $u+g+r$, and $u+g+r+i$) and the number of observations to take at high airmass per year (1 or 2 per year). Even with 2 high airmass observations per year, we would still expect some area of the sky to fall in chip and raft gaps. It is also worth noting that in our baseline simulation, we observe a spot on the sky in u typically 60 times, or 6 times per year. Taking 2 high airmass observations per year in u decreases the final coadded depth by 0.15 mags.

Figure 41 shows the science impact is fairly minimal, but we tend to lose $\sim 0.1-0.2$ magnitudes of final coadded depth.

4.17. *FBS 1.6* `even_filters`*: Sky brightness and Filter choice*

The baseline simulation is fairly aggressive in using redder filters in bright time, in order to maximize the final coadded depths. However, this can create long gaps in light curves with no bluer observations. This family of simulations removes the per-filter consideration of sky brightness (in various combinations of filters) to avoid this effect. There is a simulation where only u avoids bright time, and another with both the u and g filters avoid bright time; the simulations include both the standard baseline footprint and a variation on the 'big sky' footprint (the 'alt' runs). Figure 42 shows the resulting filter distributions in year one. Unlike the baseline simulations, there are no longer sections of several days where only y is observed.

While the goal of these simulations was to improve SNe Ia lightcurves, the gains appear to be minimal compared to the baseline strategy (Figure 43).

4.18. *FBS 1.5* `greedy_footprint`*: Always taking pairs on the ecliptic*

The survey footprint is applied to both greedy (twilight, single visits) and blob (non-twilight, pairs) visits in the scheduler, and since the final number of visits is fixed, this means that visits taken during the greedy period essentially decrease the number of visits taken during the blob period (or vice versa). This is unavoidable for most of the sky; visits taken during periods of rapidly changing conditions are not good candidates for the bigger block scheduling that occurs during blobs. However, there is some motivation for trying to take only pairs along the ecliptic, in order to maximize

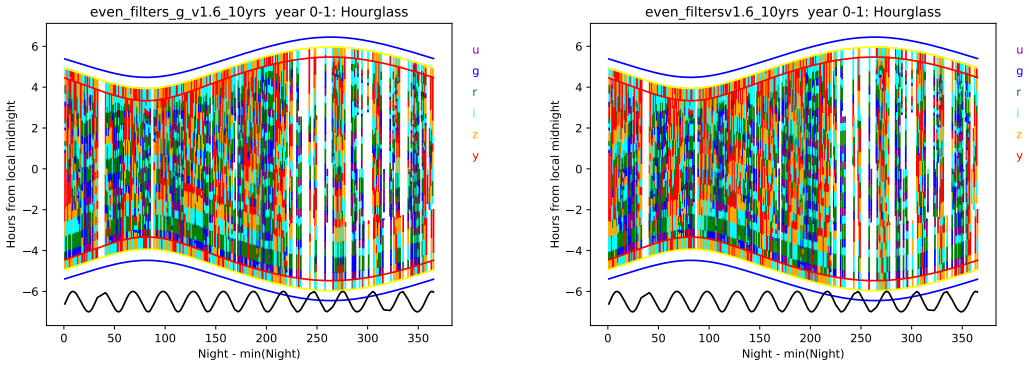

Figure 42. The filter distribution for the even filter simulations. Unlike the baseline simulations, bluer filters are observed in bright time.

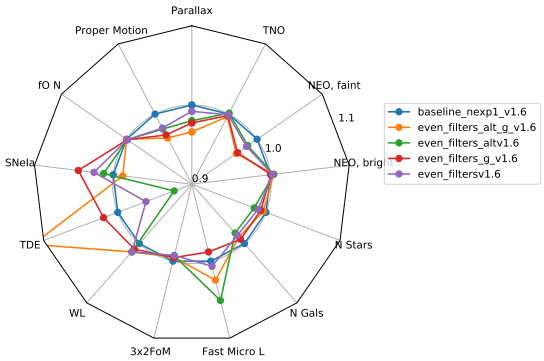

Figure 43. Science performance for the Even Filters runs. Taking bluer filters in bright time can improve SNe performance and fast transients, but is detrimental to Solar System science. The loss of depth shows up in most of the other metrics as well.

the number of visits suitable for detecting solar system objects. This simulations prohibits the greedy algorithm from observing near the ecliptic - this ensures that all observations near the ecliptic are taken in pairs.

The overall science impact is shown in Figure 44; the impact is fairly modest. The change to solar system discovery is very minimal (around a percent); this is somewhat surprising given the expected improvement with visits in pairs. Looking in depth at the number of visits acquired in pairs with *griz* filters in this simulation compared with the baseline simulation, there are more pairs along the WFD portion of the ecliptic in the greedy_footprint simulation, but fewer pairs in the rest of the WFD (see Figure 45). While more pairs on the ecliptic would help TNO discovery, it is also likely that there were 'enough' pairs there already to saturate discovery (and what is needed are deeper visits in order to discover more objects). For other populations, their on-sky distribution is wider than the strip along the ecliptic, so it may have a smaller impact on the overall discovery rate (and be partly counter-acted by the fewer visits in pairs away from the ecliptic). The better solution here is likely simply finding a good method to add pairs into the twilight observing period.

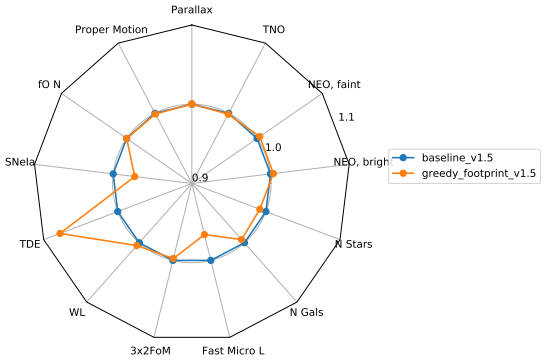

Figure 44. Science impact of not permitting greedy observations near the ecliptic.

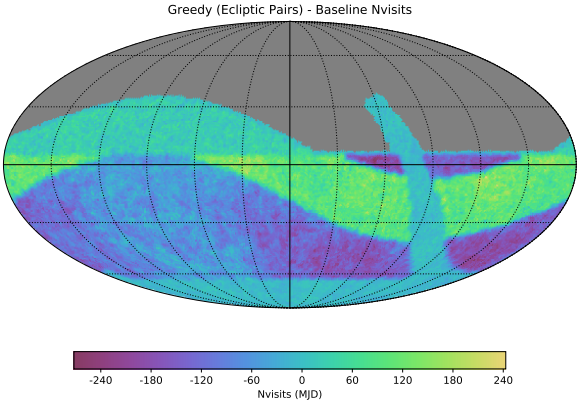

Figure 45. The change in the distribution of *griz* visits in the blob (pairs) portion of the survey, between the baseline and the greedy_footprint (pairs on ecliptic) simulation. The baseline distributes pairs of visits in *griz* (the most sensitive bands for solar system object detection) across the WFD fairly evenly; the greedy_footprint survey adds a concentration of pairs along the ecliptic within the WFD, but at the expense of pairs in the rest of the WFD.

4.19. *FBS 1.5* `spiders`*: Spider Alignment*

There are some simplifications in artifact identification as a result of difference imaging if diffraction spikes are aligned along CCD rows and columns. This experiment fixed the camera rotator so that the diffraction spikes (caused by the secondary mirror spiders) would always be aligned with the camera. This may result in the camera rotator angle being much less randomized than our baseline rotational dithering strategy; the rotation angle of the sky itself in the camera will still vary due to the alt-az telescope mount. There is little impact on our science metrics, but we note we do not currently have a metric that directly measures weak lensing systematics. See Figure 46.

4.20. *FBS 1.5: Aliasing*

One of the concerns for the survey, especially if observations were placed uniformly on the meridian, was about aliasing for periodic sources. We did not specifically

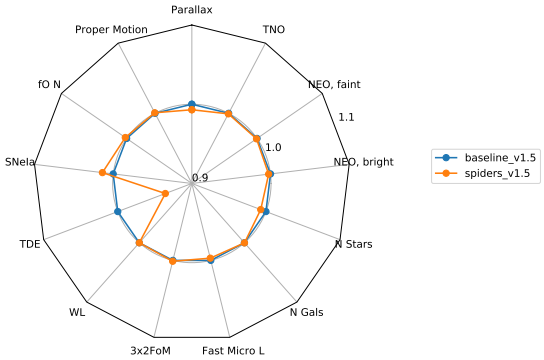

Figure 46. Science impact of keeping diffraction spikes aligned along rows and columns (Section 4.19.

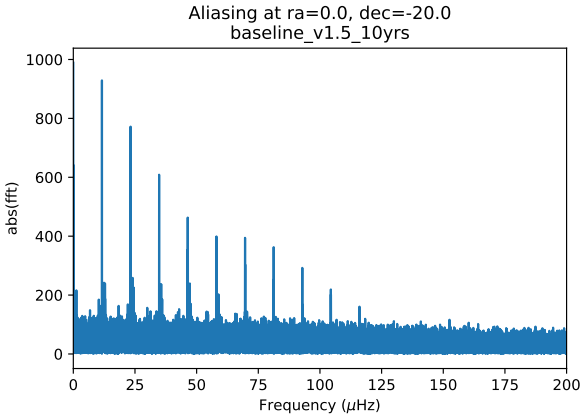

Figure 47. Aliasing at a sample position in a baseline simulation. There are peaks at harmonics of 24 hours (11.57 μHz), but this is inevitable with a ground-based telescope. The aliasing seems much lower than earlier version of OpSim where harmonic peaks could be seen past 200 μHz.

work to remove aliasing from the simulations, however we did look for the presence of aliasing in the visit cadence in these newer simulations. Figure 47 shows the FFT of observations at a sample WFD point in the baseline simulation. There is some aliasing at $\sim$ 1 day which is inevitable for any ground-based telescope. The aliasing is much lower than the `minion_1016` simulation that was analyzed in the Bell et al. cadence white paper. Attempting to plan the timing of visits to reduce aliasing is more complicated and expensive than what is currently contained in the scheduler; we therefore don't plan to add anti-aliasing features into the scheduler at this time, but to check for aliasing effects in the outputs.

4.21. *FBS 1.2: Target of Opportunity*

We performed simulations with earlier versions of the scheduler to look at the potential for following up ToO events, namely looking for the optical counterpart to gravitational wave detections. We could detect $\sim$55% of the simulated events, however, that often required pushing observations to high airmass or observing regions

outside the WFD area. Our ToO simulations used the simple followup strategy of trying to observe a target area three times in g, r, and i. As part of the final observing strategy, we should define when Rubin will attempt to observe ToOs (airmass limits, footprint limits), and what the ideal ToO discovery strategy involves regarding filter distribution and dither strategy. The short summary here is that ToOs tend to count as a separate minisurvey and do not easily just incorporate themselves within desired visits as part of the existing survey strategy (e.g., WFD); their 'cost' is the amount of time devoted to the minisurvey. We did not repeat these experiments with later versions of the scheduler; more information is required on how ToO targets would be chosen and triggered.

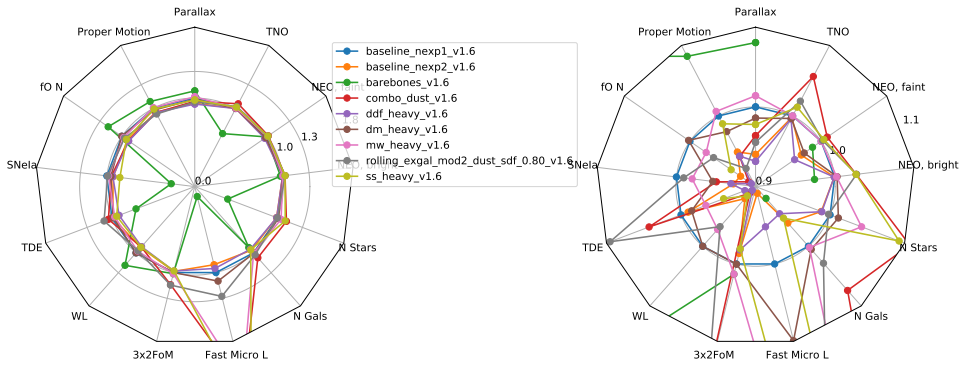

Figure 48. The science impact for the different version 1.6 simulations.

5. FBS RELEASE V1.6: CANDIDATE RELEASE RUNS

Here we describe the runs done as part of the 'candidate baselines' in the FBS 1.6 release. This set of simulations is unlike the previous experiments, in that instead of varying a particular survey strategy option across a family, we have attempted to set up a limited number of simulations that attempt to strongly boost particular goals. They are examples of more extreme choices for survey strategies; many of these options have serious drawbacks when considering an overview of science. Calling these 'candidate baselines' in no way implies they are better than some of the survey strategies explored in the FBS 1.5 release, nor that all of these would be suitable choices for an initial survey strategy; they are just intended to be exploratory examples.

The radar plots for all of these FBS 1.6 simulations are show in Figure 48.

5.1. *FBS 1.6 Baseline*

For the baseline strategy, we set the footprint to have 18,000 square degrees dedicated the the WFD survey. The WFD has a filter distribution of $u : g : r : i : z : y$ of 0.31:0.44:1.0:1.0:0.9:0.9. We include coverage of the Galactic Plane (GP) and South Celestial Pole (SCP). These areas are set to have 20% the number of counts of the WFD (if a spot in the WFD has 900 visits, points in the GP and SCP will have 180 visits). The GP and SCP are set to have equal number of visits in all filters. The North Ecliptic Spur (NES) is observed with only the g, r, i, and z filters. The NES area is set to have one-third the number of visits of the WFD. The filter distribution in the NES is set to $g : r : i : z$ of 0.2:0.46:0.46:0.4.

The total breakdown of target observing time is 85% for WFD, 6% for the NES, 6% for the GP and NES, and 5% for DDFs.

While the different survey areas are covered to different depths, the baseline scheduler treats them identically and only tries to maintain the proper ratios of area coverage. This means blocks of observations can be scheduled that cover the different regions seamlessly. It also means we have no additional constraints on how the regions are observed. For example, we currently do not reserve "good seeing" time for the WFD area.

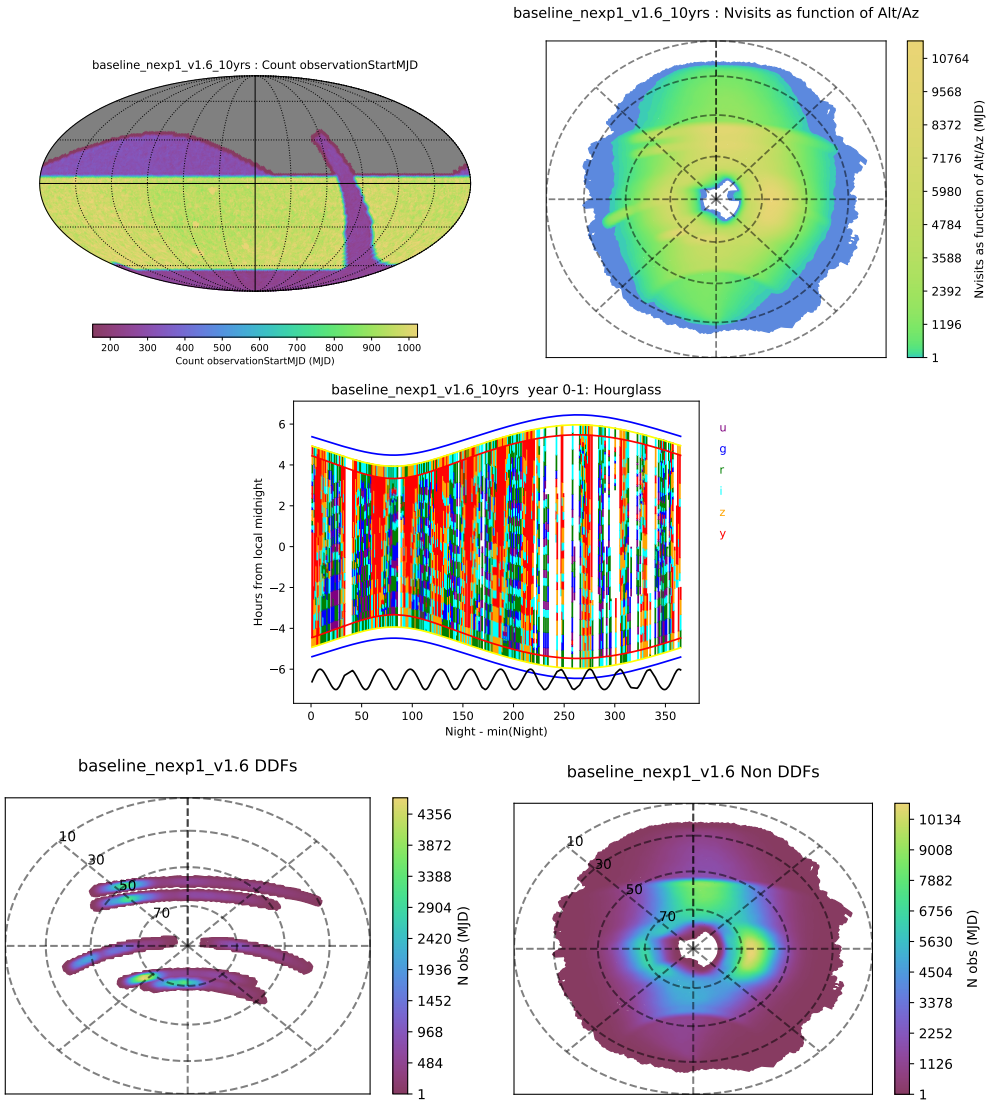

Figure 49. The baseline v1.6 simulation. The top panels show the distribution of visits (all filters) in RA/dec and Alt/Az. The middle panel shows the first year of observations color-coded by what filter was loaded. White regions represent scheduled and unscheduled downtime as well as weather downtime. The black curve on the bottom shows the moon phase. The bottom panels show the Alt/Az distribution of pointings for DDF observations (left) and non-DDF (right) on a linear stretch.

The baseline survey includes the 4 announced Deep Drilling Fields as well as a pair of fields that overlap the Euclid Deep Field South. Each individual DDF is set to take a maximum of 1% of the total visits (the Euclid pair of fields are set to a maximum of 1% combined). The standard DDF sequence is ux8, gx20, rx10, ix20, zx26, and yx20, all with 30s exposures. For any given sequence, only the five filters loaded in the camera are executed. By default, we remove the u filter when the moon is more than 40% illuminated at the start of the night.

We run 2 baseline simulations, one with 1x30s visits and one with 2x15s visits. The main difference is the additional readout time in the 2x15s version drops the open

shutter fraction from 77% to 72%. This puts the 2x15s simulation close to failing the SRD FO metric, with some parts of the WFD region only reaching 824 observations (the median is still 892).

For the rest of the simulations in v1.6 we use 1x30s visits. If 2x15s visits are required there will be a significant drop in the number of visits, and areas outside of the WFD may need to be scaled back to still meet SRD requirements.

When it is non-twilight time and we are not observing DDFs, we use a Markov Decision Process to dynamically build a queue of observations. Observations are planned in 44 minute blocks (22 minutes for an initial area, 22 minutes to repeat the area). The size of the blocks can scale slightly to try and fill time before twilight (e.g., it will expand to a pair gap of 25 minutes if there are 50 minutes until morning twilight begins). All observations are taken in pairs, with potential combinations of $u+g$, $u+r$, $g+r$, $r+i$, $i+z$, $z+y$, or $y+y$. The ordering of the filter pairs can change depending on what filter is currently loaded (e.g., if the scheduler decides to observe a $g+r$ sequence, the r observations will be taken first to eliminate a filter change if possible.)

The camera rotator angle (relative to the telescope) is randomly set each night between -80 and 80 degrees. The angle is set when the block is scheduled, so there can be a few degrees of drift between when the rotator angle is computed and when the observation is actually taken.

The MDP uses basis functions based on

- The 5-sigma depth (for both filters in the pair being taken)
- The footprint uniformity (again, in both filters)
- The slewtime
- Staying in the current filter
- Rewards taking 3 observations per year per filter over the entire survey footprint

The MDP also includes basis functions that are simple masks

- Zenith is masked (to avoid long azimuth slews)
- 30 degrees around the moon is masked
- The bright planets (Venus, Mars, and Jupiter) are masked with a 3.5 degree radius

If the sun is higher than -18 degrees altitude, or there is not enough time remaining to take observations in pairs, the scheduler reverts to a greedy algorithm and selects observations one at a time. We use a similar MDP for these greedy twilight observation decisions.

Compared to many of the other FBS 1.6 candidate baseline simulations, the baseline spends a lot of time observing the WFD, with a median of 948 visits. The higher

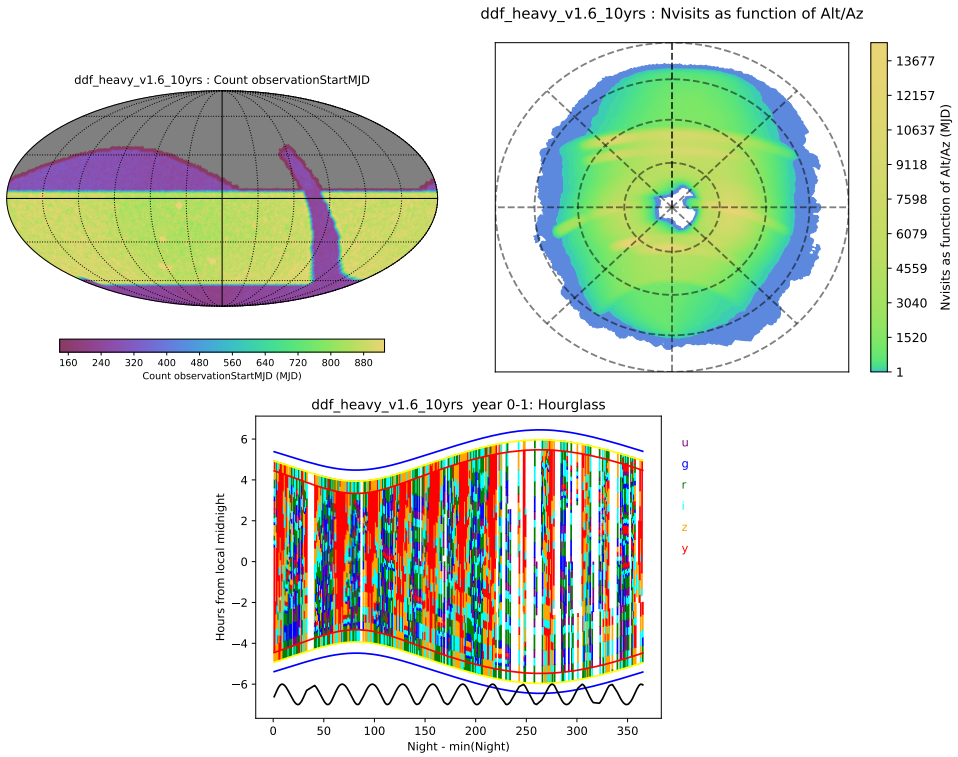

Figure 50. DDF Heavy simulation. Nearly identical to the baseline, but giving as much time as possible to DDF observations.

number of visits means a faster cadence and better sampled lightcurves for objects with durations comparable to a season length. Our baseline simulation also has very light coverage of the Galactic bulge, resulting in fewer fast microlensing events than other potential footprints.

5.2. *DDF Heavy*

This run is nearly identical to the baseline, but gives a large fraction of time to the deep drilling fields. Each of the five DDFs takes between 2.4 and 2.9% of the survey, with 13.4% of all visits being used for DDF observations. The baseline has 4.6% of visits used for DDFs. This is enough time that the WFD area near the DDFs fails to reach 825 visits over 10 years, but the SRD requirement is formally still met because the median WFD point is observed 875 times.

As expected, the majority of non-DDF science cases suffer if we dedicate such a large fraction of time to the DDFs. It is worth noting that most metrics within MAF are not tailored for DDF purposes; this is an area that is missing science metrics.

5.3. *Barebones*

The barebones simulation is not a viable survey strategy, but provides an extreme example where we focus exclusively on meeting the SRD requirements, with little optimization for science.

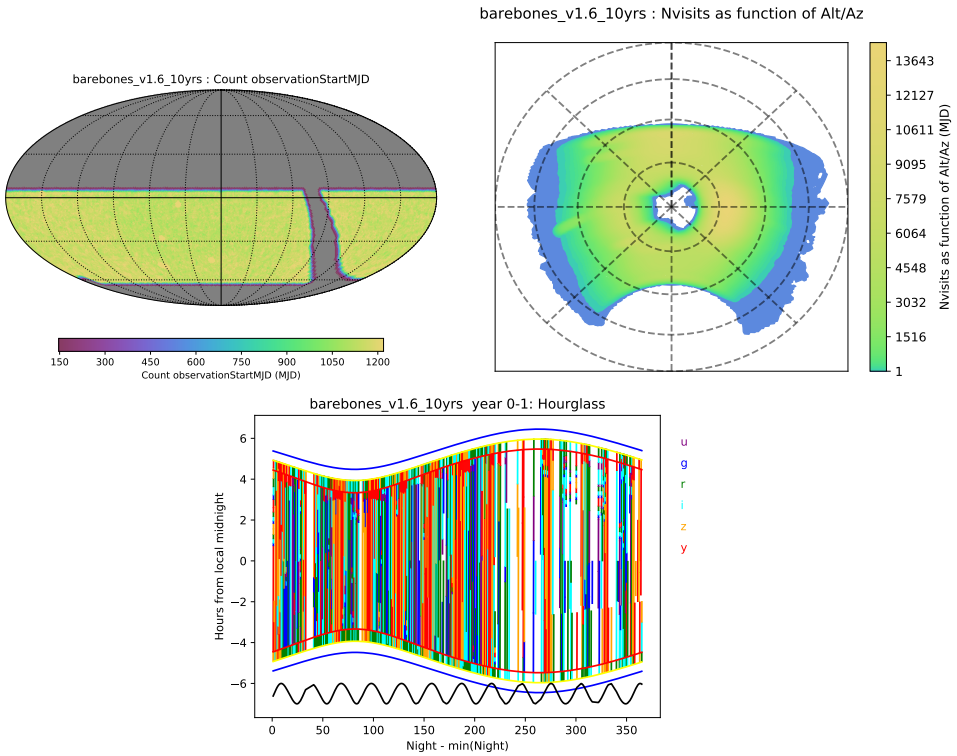

Figure 51. The barebones simulation covering just the WFD area as efficiently and deeply as possible.

The survey footprint is restricted to the baseline 18,000 square degree WFD area only. Deep drilling fields are included, but capped at $\sim 2.5\%$ of the total visits. Visits in u and y are unpaired, while the rest of the filters are paired in the same filter. This results in very few filter changes in a night.

There are a wide number of reasons why this would be a terrible survey strategy – detected transients would have no color information, photometric uber-calibration could be difficult with the galactic plane gap, a lack of solar system object because the NES is not included, etc. The main purpose is to show the scheduler can run very near the theoretical maximum for open shutter fraction, with this run reaching 80%. Also, we can note the fONv MedianNvisits metric reaches 1155, which is 40% higher than the SRD requirement of 825. This also implies that we can observe a maximum of ~ 115 WFD visits per year in the event we want to adjust the scheduler to attempt to catch up on the WFD progress.

The total lack of bulge coverage means the barebones simulation contains virtually no fast microlensing events. Taking pairs in the same filter also radically reduces the number of SNe Ia that are well measured.

5.4. *Data Management Heavy*

This simulation is similar to the baseline, but includes various modifications that may be helpful for Data Management purposes. Across the WFD region, u, g, and r a few images per year are taken at high airmass so that DCR correction models can be

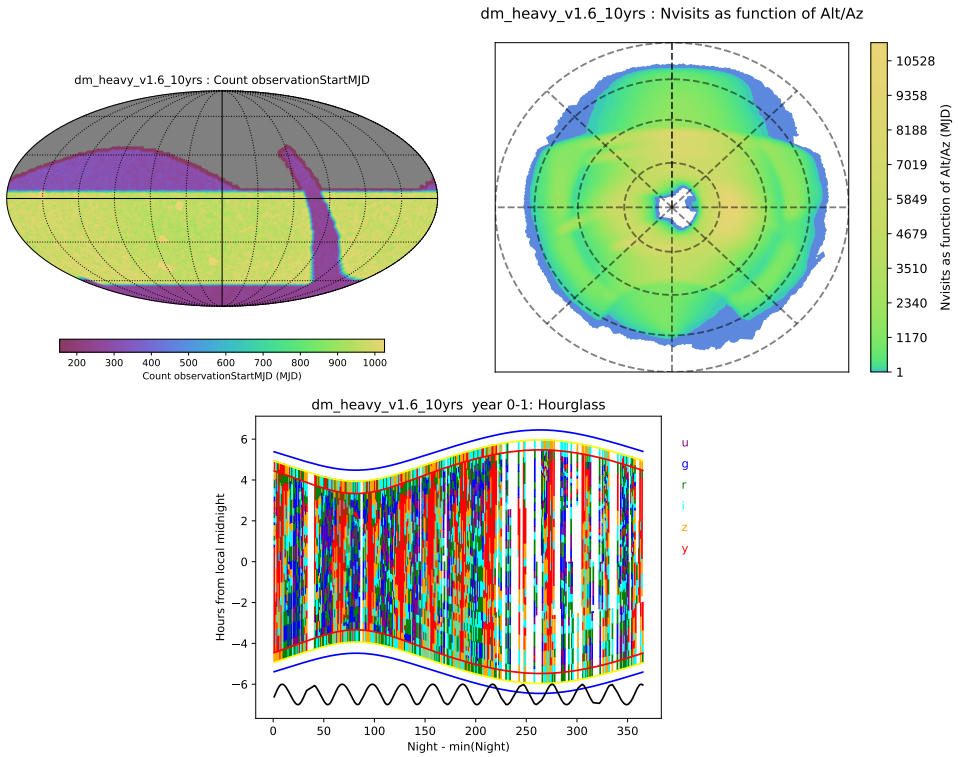

Figure 52. The DM heavy simulation. Similar to the baseline, but the alt/az plot shows how some observations are being taken at high airmass to support DCR modeling.

made. The camera rotator angle is set so that diffraction spikes fall along CCD rows and columns. This helps with difference imaging so the maximum possible area can be used, but may result in weak lensing systematics. Each year, the scheduler prioritizes taking g, r, and i images of the whole sky in good seeing conditions (defined as 0.7″effective FWHM or better). The DDF fields use larger dithers, up to 1.5 degrees, compared to the default 0.7 degree maximum.

The addition of images taken at high airmass has a small negative impact on most science cases.

5.5. *Rolling Extragalactic*

The rolling extragalactic is motivated by cosmological drivers. The footprint is modified so the 18,000 square degrees of the WFD are placed in low-extinction regions. The simulation also executes a half-sky rolling scheme, which should result in better sampled lightcurves for extragalactic transients.

This simulation divides the sky into quarters, and has one northern stripe and one southerns stripe with a rolling emphasis at a time. This could be preferable to a simple two-band rolling scheme, because with the quarters a region of emphasis will always be available to northern telescopes. If we rolled with an emphasis purely on the southern half of the WFD region, ∼80% of the Rubin alert stream would become unavailable to northern hemisphere observatories for that season.

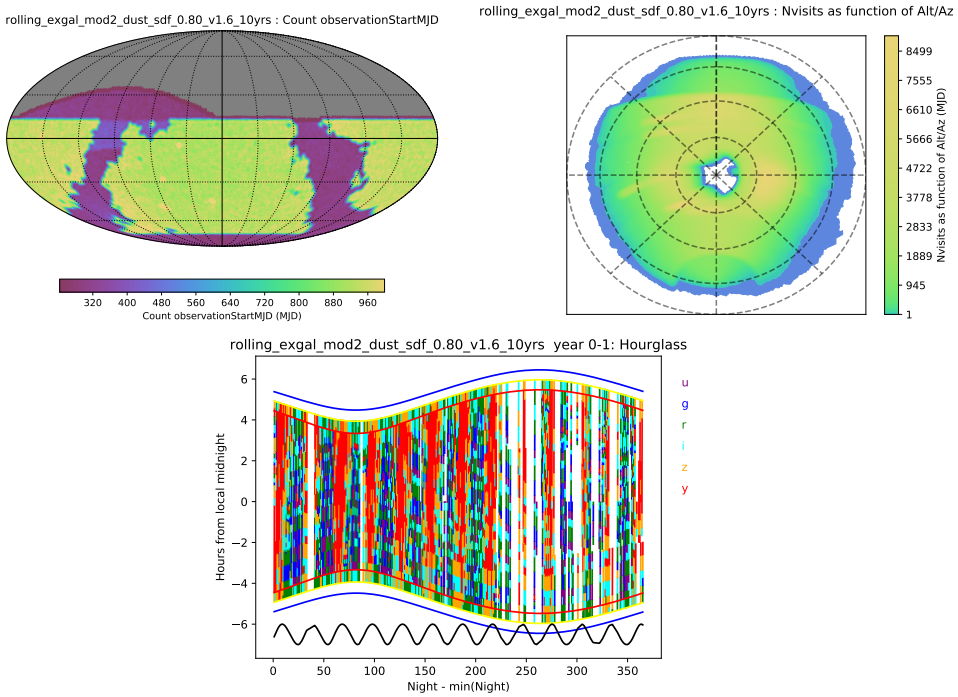

Figure 53. The Rolling Exgal simulation. The WFD area is set to be 18,000 square degrees of low extinction area.

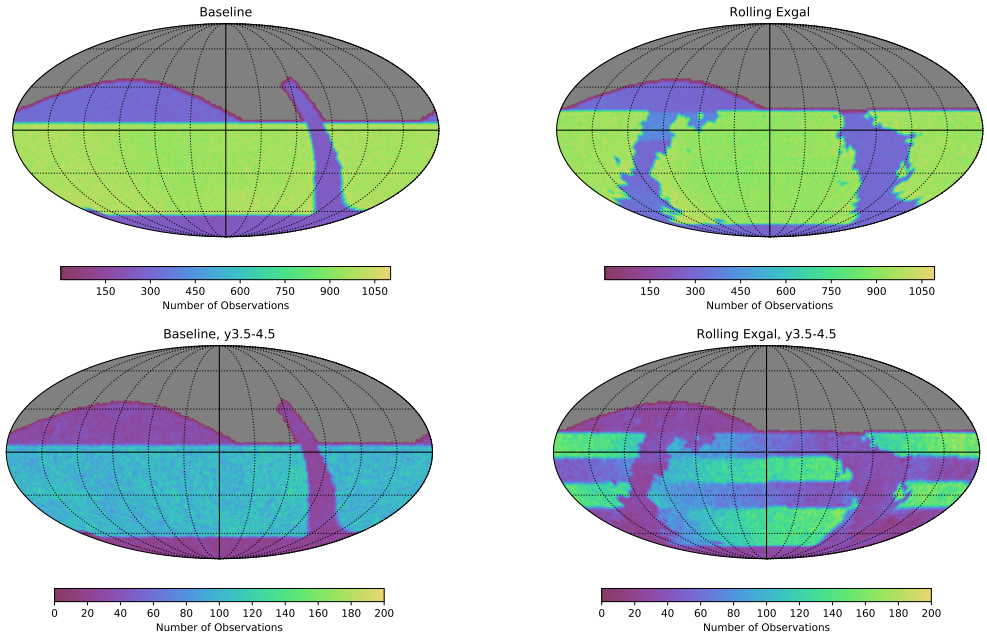

Figure 54. Illustration of rolling cadence. The top panels show the number of observations after 10 years (all filters) for the Baseline and Rolling Exgal simulations (excluding DDF observations). Both simulations have very smooth WFD coverage, with ∼900 observations. The lower panels show the number of observations taken between 3.5 and 4.5 years. The baseline WFD remains smooth, while the Rolling Exgal simulation has declination stripes of high and low counts.

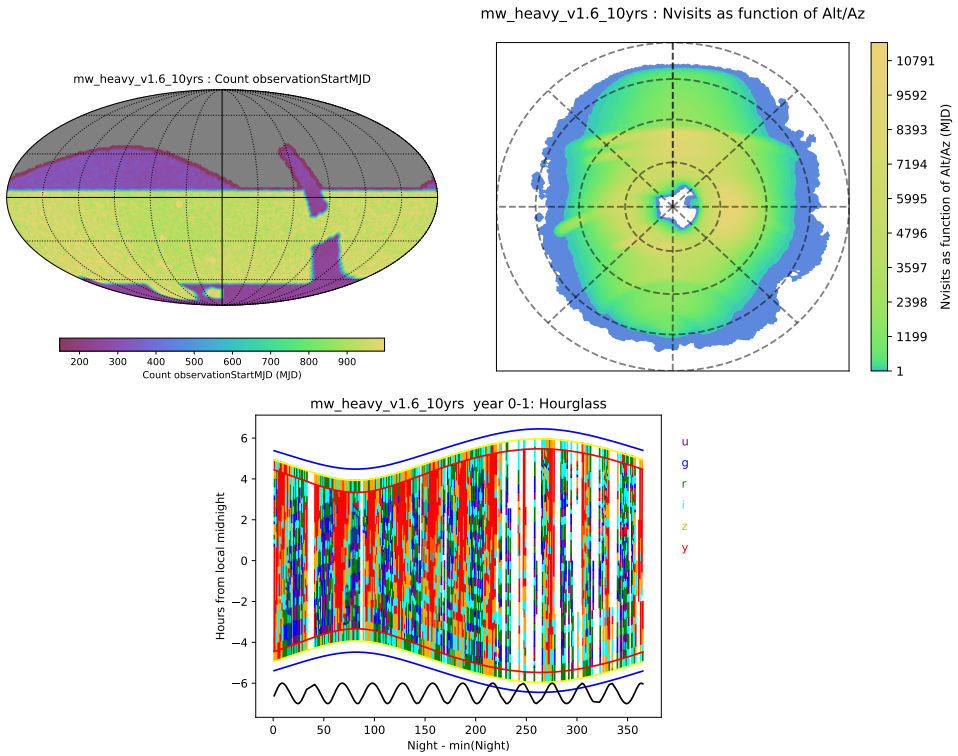

Figure 55. The Milky Way heavy simulation. Similar to the Baseline, but the bulge and Magellanic Clouds are added to the WFD area.

As expected, avoiding high extinction regions increases the number of galaxies. We expect the addition of rolling will show improvements in more sophisticated SNe Ia metrics from the community. The footprint covers some of the Magellanic Clouds, boosting the fast microlensing events. The science gains come at the expense of some of the SRD metrics.

5.6. *Milky Way Heavy*

The Milky Way heavy simulation covers the Galactic bulge, LMC, and SMC as part of the WFD area.

There is very little change in the overall median coadded depths compared with the baseline since the extra WFD area is added to a region of the sky that is under-subscribed in the baseline. In the baseline simulation, there are an excess of observations in the WFD on either side the galactic plane, so covering the bulge is "free", in the sense that it uses these excess pointings to cover the bulge.

There is a large boost in microlensing events and number of stars, with little impact on the other metrics. We would benefit from other metrics for bulge-specific science cases to explore using a different filter distribution for the bulge region.

5.7. *Solar System Heavy*

For the Solar System Heavy simulation, the baseline survey footprint is modified to include ecliptic plane coverage through the galactic plane.

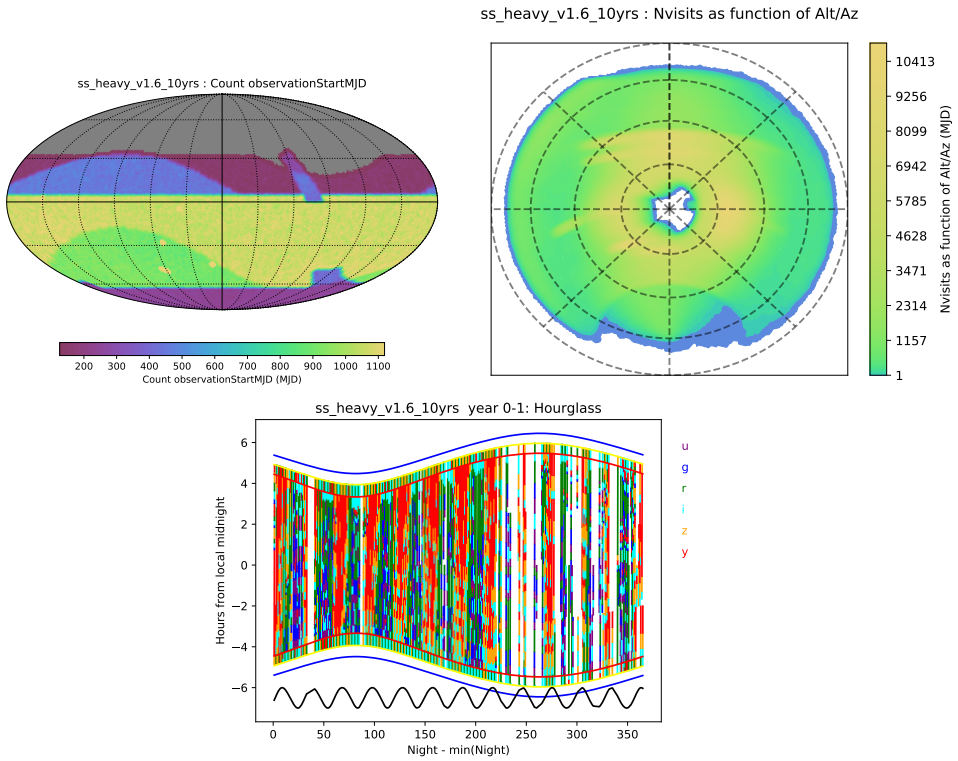

Figure 56. The Solar System heavy simulation. The high airmass observations are twilight NEO observations.

A fraction of twilight time is used for a NEO survey in r band. The NEO survey uses very short (1 second) exposures at high airmass (toward the sun in evening or morning twilight). This simulation only uses i, z, y in the remainder of the twilight time, making sure we observe more r-band in non-twilight and in pairs. It also includes $r+r$ pairs in non-twilight time. For regular 1x30s visit twilight observations, we avoid observing the ecliptic, thereby ensuring they are always taken in pairs in non-twilight time.

The simulation shows a slight improvement in the discovery of bright NEOs and TNOs, with a slight decrease in discovery of faint objects of all populations, while significantly impacting SNe Ia discovery due to the addition of pairs in the same filter. Solar system metrics, particularly for bright objects, are most sensitive to footprint (they tend to get enough visits to discover objects, so need to explore more area of sky) so a larger footprint (such as the big sky style with galactic plane and extended northern sky coverage) works better. For fainter objects, visits in redder filters and in the same filter are ideal; beyond that, more visits are important as the timing of discovery is more critical. The Solar System Heavy simulation adds enough twilight visits that the overall number of long exposure ($>$ 1s) visits in the traditional WFD footprint is reduced by about 6%; this has an impact on the detection of faint objects. Further optimization toward solar system objects would likely produce

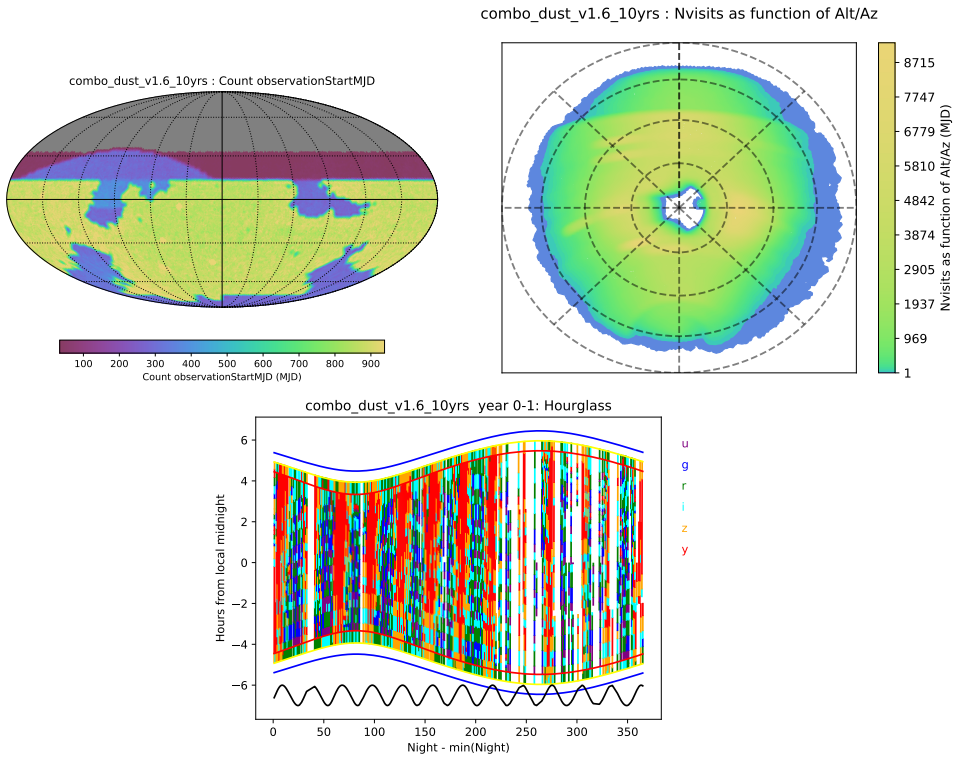

Figure 57. The Combo Dust simulation. Similar to the Rolling Exgal simulation, but the WFD is expanded to include the bulge and ecliptic, Magellanic Clouds, and an anti-center bridge.

a slightly different simulation to this, however this is a reasonable example of the trades with other science.

5.8. *Combo Dust*

This simulation attempts to improve several science cases compared to the baseline simultaneously. The footprint used here starts with defining the WFD area as 18,000 square degrees with low extinction. Then an additional 2,000 square degrees are added to WFD to cover the bulge, the ecliptic through the galactic plane, the LMC and SMC, and an outer Galactic plane region. Dusty areas of the sky and the South Celestial Pole are covered at about one-quarter the WFD depth. The NES is covered in g, r, i, and z. The footprint also includes very light coverage to the northern limit of the telescope in g, r, and i so there can be templates for ToO events on the entire accessible sky. This simulation includes the same half-footprint rolling scheme as Rolling Extragalactic.

The footprint has 35 free parameters for setting the various region locations and filter ratios. Many of these have have been set by eye or use historical values; it is quite likely these parameters could be improved.

This simulation manages to boost nearly all the science metrics at the expense of reducing margin in the SRD metrics. When we run the *combo_dust* with 2x15s visits, the fO metric drops below the SRD requirement of 825 visits to 817 visits. The

filter	Baseline	Baseline 2 snaps	Barebones	DDF Heavy	DM Heavy	MW Heavy	Rolling Exgal	SS Heavy	Combo Dust
	(mags)	$m_{\rm Baseline} - m_{\rm Sim}$							
u	25.86	0.24	-0.13	0.08	0.11	0.02	0.11	-0.02	0.12
g	26.97	0.11	-0.15	0.09	0.12	0.01	0.10	0.07	0.14
r	26.95	0.08	-0.12	0.08	0.07	0.01	0.10	0.05	0.14
i	26.40	0.07	-0.17	0.11	-0.01	0.01	0.11	0.11	0.15
z	25.67	0.06	-0.12	0.08	-0.01	0.01	0.11	0.02	0.11
y	24.90	0.06	-0.14	0.06	0.04	0.01	0.09	0.03	0.09

Table 4. Difference in median coadded five sigma depths compared to Baseline for v1.6 simulations. Negative values indicate deeper depths.

footprint can be adjusted to meet the SRD requirement, but it does imply there will be very little contingency if we use 2-snap visits. The 1x30s visit *combo_dust* has a median of 885 visits in the WFD region, meeting SRD requirements.

6. CROSS-FAMILY SURVEY STRATEGY CHOICES

The families of simulations described in the previous sections (FBS 1.5 and 1.6 releases) maintained the approach of varying a single kind of parameter (such as the amount of time devoted to triplets of visits in the `third_visit` runs), but the underlying survey strategy choices can cover multiple families (such as the runs including pairs in mixed or the same filters and the triplet visits runs). In this section, we will consolidate and summarize some of the results from the previous sections if they cut across families (presenting some additional metric results in the process). Strategy questions that are confined to a single family are not repeated in this section.

6.1. *Visit Exposure Time*

There are several choices related to the individual visit exposure time:

- Should visits be made of a single exposure (1x30s) or two (2x15s)?
- Should u band visits be longer exposures (60s)?
- Should the visit exposure time be variable?

For the first of these questions – 1x30s visits or 2x15s visits – is perhaps the most important, yet the final decision cannot be made until the camera is on the sky and acquiring data, to verify cosmic ray rejection will work and that the two snaps are not required. In addition, there are still some rapid transient or variable science cases that may still benefit from 2x15s visits. However, this decision does weigh in on other survey strategy choices. The decrease in overhead going from 2x15s visits to 1x30s visits allows 8% more visits to be obtained in the same time, which allows significantly more options for footprint coverage and addition of minisurveys.

The second question – increasing the u band exposure time – was addressed in Section 4.14. The tension comes between increasing coadded depth in u versus a dramatically decreased number of u band visits making it more difficult to obtain u photometry for transients. Extending the visit exposure time while also maintaining the number of u band visits would require adding about 6% more visits to WFD by removing them from other minisurveys or other filters (note that shifting visits into bluer filters overlaps with some of the options discussed in Section 4.7). We should run additional simulations to evaluate this impact, using an updated survey footprint and choice of rolling cadence.

The third question – variable exposure time – was addressed in Section 4.15. Adding a variable exposure time decreases the scatter in the individual image five sigma limiting magnitudes, but results in fewer overall visits. It had mixed effects on the typical limiting magnitudes per visit, and the coadded u band depth became shallower by 0.3 magnitudes, although there are many parameters which could be tuned to attempt to improve this (making the goal u band visit depth deeper could be one, although it is complicated because of the u band read-noise limitation). In general,

this seems to add complication to practical aspects of the scheduler when using real-world telemetry (due to the difficulty in predicting the proper exposure time when scheduling blocks of blobs) and may not have clear advantages to science.

Figure 58 shows the metrics included in the radar plots, together with additional metrics, for all of the simulations which relate to the visit exposure time.

6.2. *Intra-night Cadence*

Within each night, there are some choices about how to acquire visits during the non-twilight, good weather time:

- Should pairs of visits be taken in the same filter or different filters?
- Should additional visits (triplets) be acquired for some portion of the fields during the night?

Whether or not pairs should be taken in the same filter or different filters presupposes that visits should be taken in pairs at all. With the current baseline software and criteria for linking moving objects, the need for pairs of visits is clear. Even without the requirement of linking moving objects, there will be many unknown solar system objects in every image; taking pairs of visits is a reasonable way to quickly filter these detections out of the alert stream.

Then the question is whether pairs should be taken in the same filter or in different filters. The simulations in Section 4.2 show that same-filter pairs achieve about 4% more visits over the survey than split-pairs, due to fewer filter changes. Same-filter pairs have a slight benefit to discovery for all solar system objects (2-3%), but same-filter pairs have a big negative impact on transients (10-50% depending on the population and metric requirements).

The final question is if some fraction of visits should be acquired in triplets. Section 4.3 explores adding a third observation per night for some fraction of the available time. By adding a third visit, less sky is covered per night. This has a small negative impact on most solar system object discovery, with NEO completeness at $H \leq 22$ dropping by about 4% for the simulation with 120 minutes per night spent on triplets. Surprisingly, the transient metrics did not show much improvement with the addition of triplets; detection of TDEs dropped by 8% with 120 minutes of triplets. Some of this is due to the existence of more than two visits per night over some fraction of the sky even in the baseline simulations, due to field overlaps. This is likely also due to the timescales over which the objects change, the smaller amount of area covered per night, and the specific detection requirements in the metrics.

It seems likely that we are also missing metrics; metrics reflecting classification confusion and detection requirements for transients which were the target of the Bianco et al white paper ('Presto-Color') may change this evaluation. Figure 59 shows the response of the radar plot metrics, together with some additional metrics, for these intra-night cadence simulations. Note that the triplets options have not been

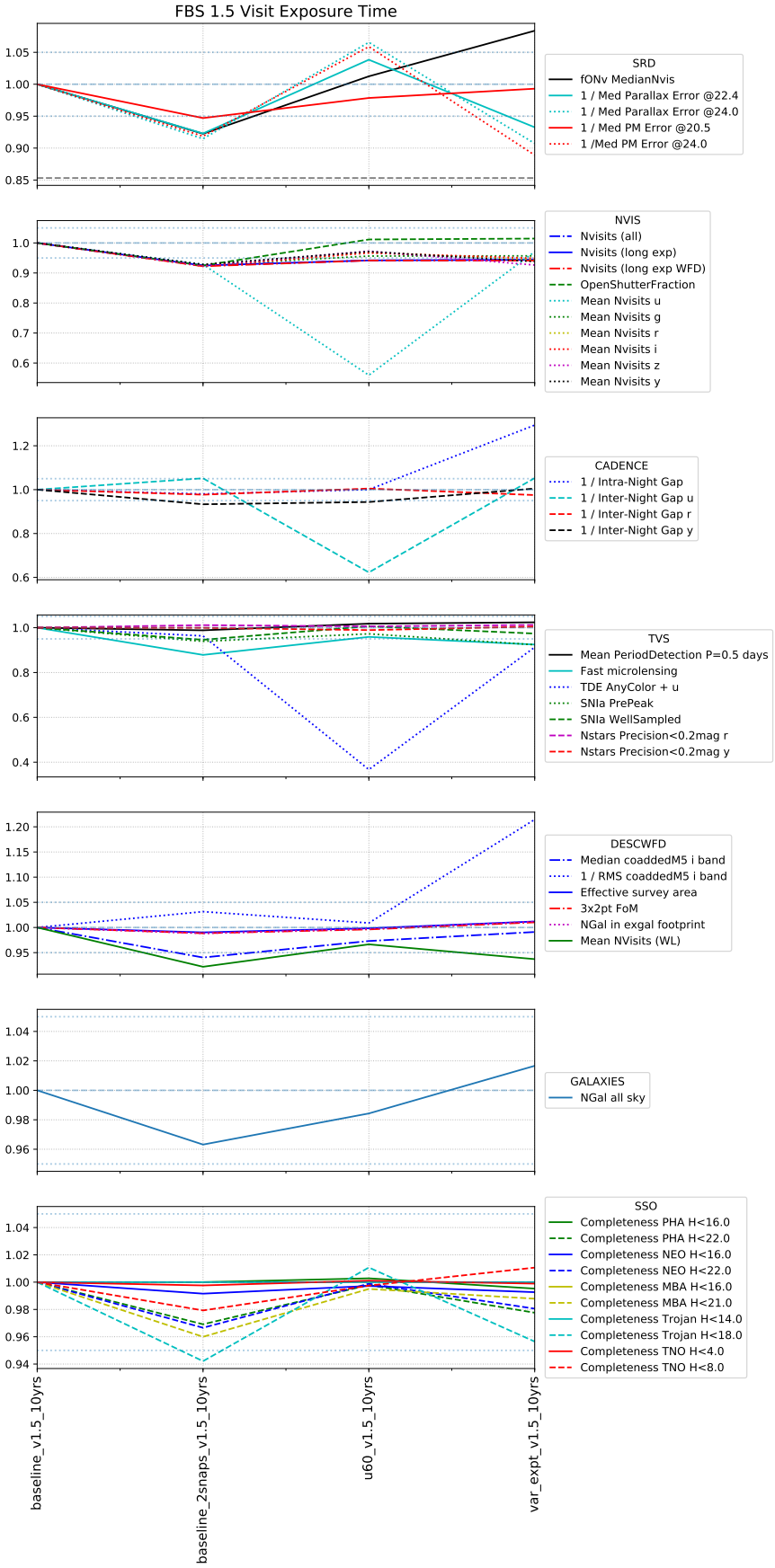

Figure 58. Comparing the simulations relevant to the visit exposure time (Section 6.1). All metrics are scaled so that bigger is better and all are normalized against the baseline run. The dashed black line on the SRD metrics subplot indicates where fONv will fail to meet requirements.

tested with survey footprints beyond the baseline or with a rolling cadence; additional simulations using an updated survey footprint and choice of rolling cadence would be useful.

6.3. *Survey Footprint*

The choice of the survey footprint is extremely significant for science. It is hard to separate the choice of the WFD footprint from other choices about footprint coverage; if the WFD footprint is moved, then the amount of sky required to be covered in the NES, GP and SCP minisurveys changes and their time requirements change accordingly. As a result, it seems reasonable to make a joint consideration of the simulations which vary over all of this coverage. We include the simulations from Sections 4.5 (4 potential WFD footprints and various minisurveys), 4.6 (a 'big sky' WFD background and variable coverage across the galactic plane), and 4.7 (where `filterdist_indx2_v1.5_10yrs` has a standard filter distribution with a limited WFD and GP only footprint).

The desired footprint options include:

- Extragalactic science needs 18,000 square degrees (the WFD) to include only low dust extinction regions, which requires rearranging the traditional footprint. Moving the declination limits of the survey north and south, then adding a dust extinction limit around the (entire) galactic plane will satisfy this requirement. It reduces the area that needs to be observed in the NES and the SCP, but increases the area needed for galactic plane studies.

- Solar system science needs coverage in the north, specifically in the region of the ecliptic plane; adding a survey extension with a moderate number of visits in the NES will satisfy this requirement.

- Transient and variable science, as well as Milky Way science, needs relatively dense coverage through the galactic bulge and moderate coverage through the rest of the galactic plane. Adding a WFD-level extension through the galactic bulge and low density coverage through the rest of the GP would likely satisfy this requirement.

- Transient and variable science also desires relatively dense coverage of the Magellanic Clouds and moderate coverage of the SCP. Dense coverage of the LMC/SMC and low density coverage of the rest of the SCP would likely satisfy this requirement.

- Overlap with Euclid and other surveys drives a desire to get low density coverage throughout a northern extension of the sky, up to about the top of the NES (in the region at higher ecliptic latitudes than the NES).

As shown in Section 4.4, the minimum amount of time needed to meet SRD requirements in the WFD is approximately 70% of the total available. Near 70%, every field

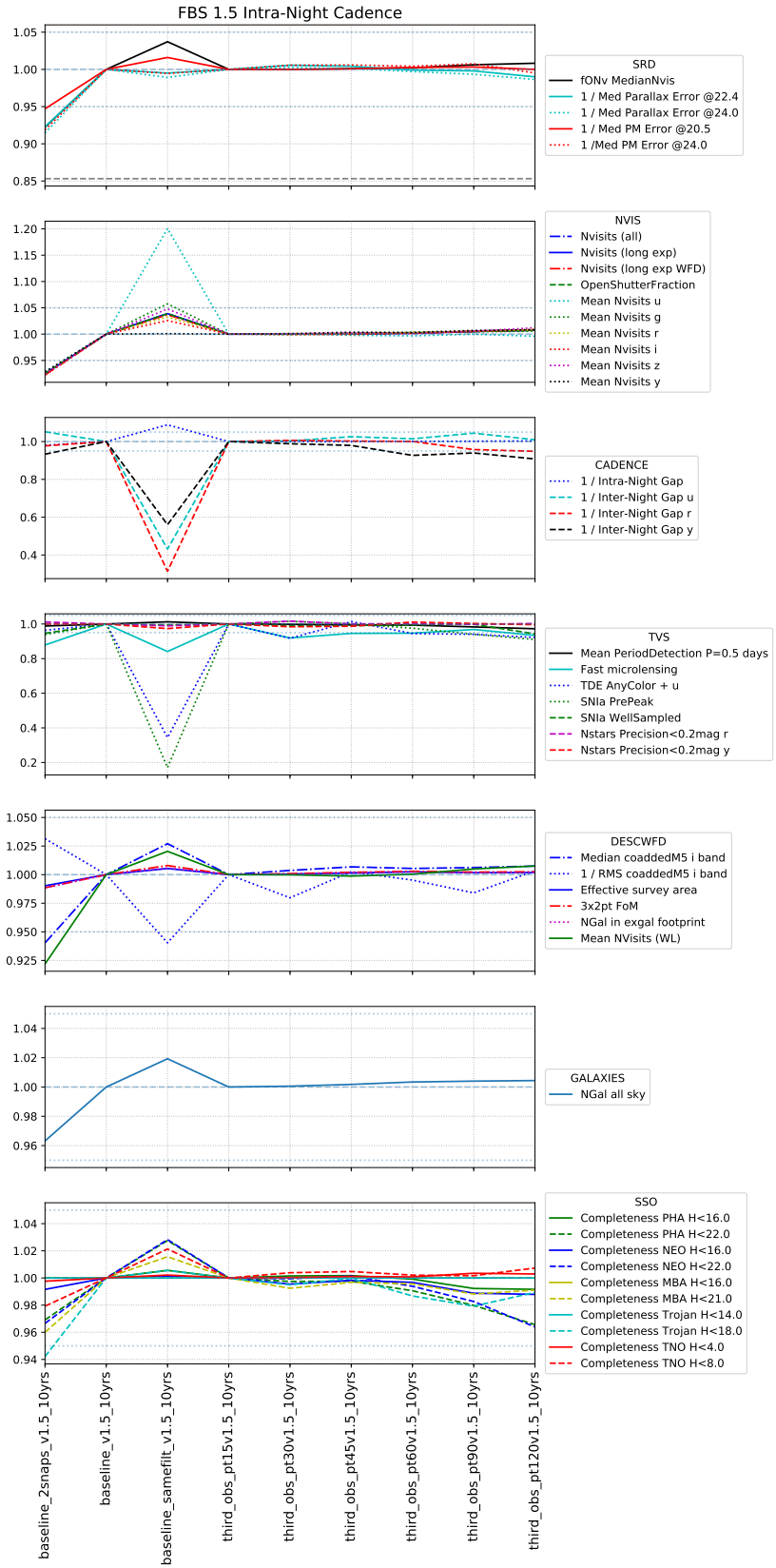

Figure 59. Comparing the simulations relevant to the intra-night cadence (Section 6.2). All metrics are scaled so that bigger is better and all are normalized against the baseline run. The dashed black line on the SRD metrics subplot indicates where fONv will fail to meet requirements.

in the WFD can receive at least 825 visits; because the WFD is so large, attempts to increase the number of visits per field require significant additions of time. To add about 50 additional pointings (5% of 825) to each WFD field requires about an additional 5% more time spent on the WFD. However, this is not exact, as can be seen when comparing the fraction of visits covering the WFD (Figure 60) and the SRD metrics at the top of Figure 61. The DDFs will require at least 5% of the survey time.

The survey footprints included in these simulations can be gathered into some broad groups (reflected in the shading in Figure 60 and 61). These groups are:

- A bare-bones, declination limited stripe across the sky (`filter_dist`). This footprint loses 30% of the TNO population, and will have other problems outlined in the Schwamb et al. whitepaper ('A Northern Ecliptic Survey for Solar System Science'). It also does not meet WFD goals for extragalactic science, due to the levels of dust extinction within the footprint.
- Variations on the current baseline WFD where the 18,000 square degree WFD includes regions of high dust extinction ('standard plus').
- A bare-bones versions of the base big sky WFD where the 18,000 square degree WFD footprint is extended N/S and is entirely low dust extinction. This footprint has no coverage of the GP or SCP, which results in losing almost all fast microlensing events as well as having a dramatic impact on the number of stars detected by the survey.
- Variations on the big sky WFD where the NES, GP and SCP are covered as well ('big sky plus').

While the bare-bones versions of the footprint which contains only WFD are likely unworkable, the two groups with variations on these footprints adding additional minisurveys are useful to compare further. In general, the 'standard plus' group puts more visits into the WFD region than the 'big sky plus' group, and this is reflected in the SRD metrics. The SRD metrics pass minimum in all runs except the `footprint_big_wfd` and `footprint_newA` runs, but the other 'big sky plus' are only about 20-40 visits (2.5-5%) per pointing above the threshold. Looking at science metrics beyond the SRD, the 'big sky plus' runs perform better than the 'standard plus' in many categories.

Some of the simulations within this group don't significantly vary the area or number of visits within the survey footprint, but vary the filter distribution. To evaluate the effects of the filter distribution more directly, it is better to start with the set of runs in the `filter_dist` group (Section 4.7. Figure 62 shows these same metrics scaled across that subset of simulations, using a run with the same footprint but the baseline filter distribution as the comparison point. The metrics from this set show the same general result as mentioned in Section 4.7: solar system science prefers redder filters

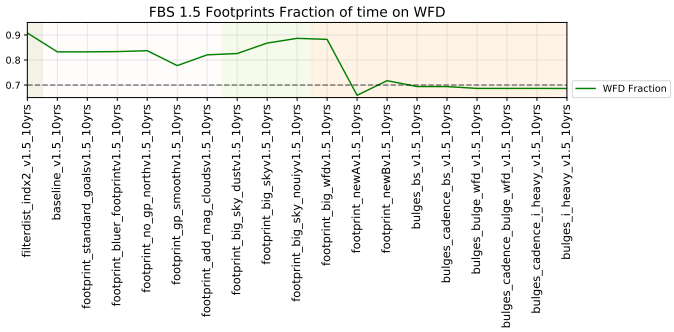

Figure 60. Comparing the fraction of visits used to observe the WFD and the total number of visits gives an appropriate fraction of time spent on the WFD. As the survey footprint changes, this fraction changes significantly. Adding galactic plane, NES, and SCP coverage (as in footprint_newA and further to the right on the plot) to the 'big sky' WFD footprint with varying levels of coverage requires a fairly large fraction of time. Similar levels of coverage for the current baseline WFD footprint require less time outside of WFD, as the galactic plane region is smaller and the low-coverage northern stripe in the 'big sky' footprints is not included. The shading groups similar kinds of footprints and matches Figure 61.

(but appreciates g band visits), the number of galaxies increases with redder filter distributions, and DESC WFD metrics (as they are based on i band coadded depth) prefer slightly redder filter distributions. Transient detection (for metrics requiring colors pre-peak in particular) prefer bluer filter distributions. The primary levers are the number of visits in u band and i band.

6.4. *Other minisurveys*

None of the survey footprint simulations above included additional minisurveys beyond DDFs, NES, SCP and GP and a small northern declination band. Simulations adding specific minisurveys are contained in Sections 4.12 (the twilight NEO minisurvey), 4.13 (the short exposures minisurvey) and 4.16 (the DCR minisurvey). Each of these minisurveys has various science advantages and disadvantages presented in the previous sections, and each requires an additional investment of time.

Figure 63 evaluates the number of WFD visits that are *not* part of the minisurvey or in the NES, SCP, GP or other footprint regions. For minisurveys where the visits are not useful for WFD visits (such as the twilight NEO survey and the short exposures minisurvey), the change in the number of WFD visits translates very simply to the amount of time required for the minisurvey. For the other minisurveys, such as the DCR minisurvey, the visits are useful for the WFD and can be counted as part of the WFD, however they are sub-optimal visits. For these kinds of minisurveys, the impact on science will vary depending on exactly how 'suboptimal' the visits are. Looking at Figure 64 for the science impacts of these minisurveys, it is clear that the impact of the DCR minisurvey is much less than the decrease in standard WFD, non-high-airmass visits would suggest.

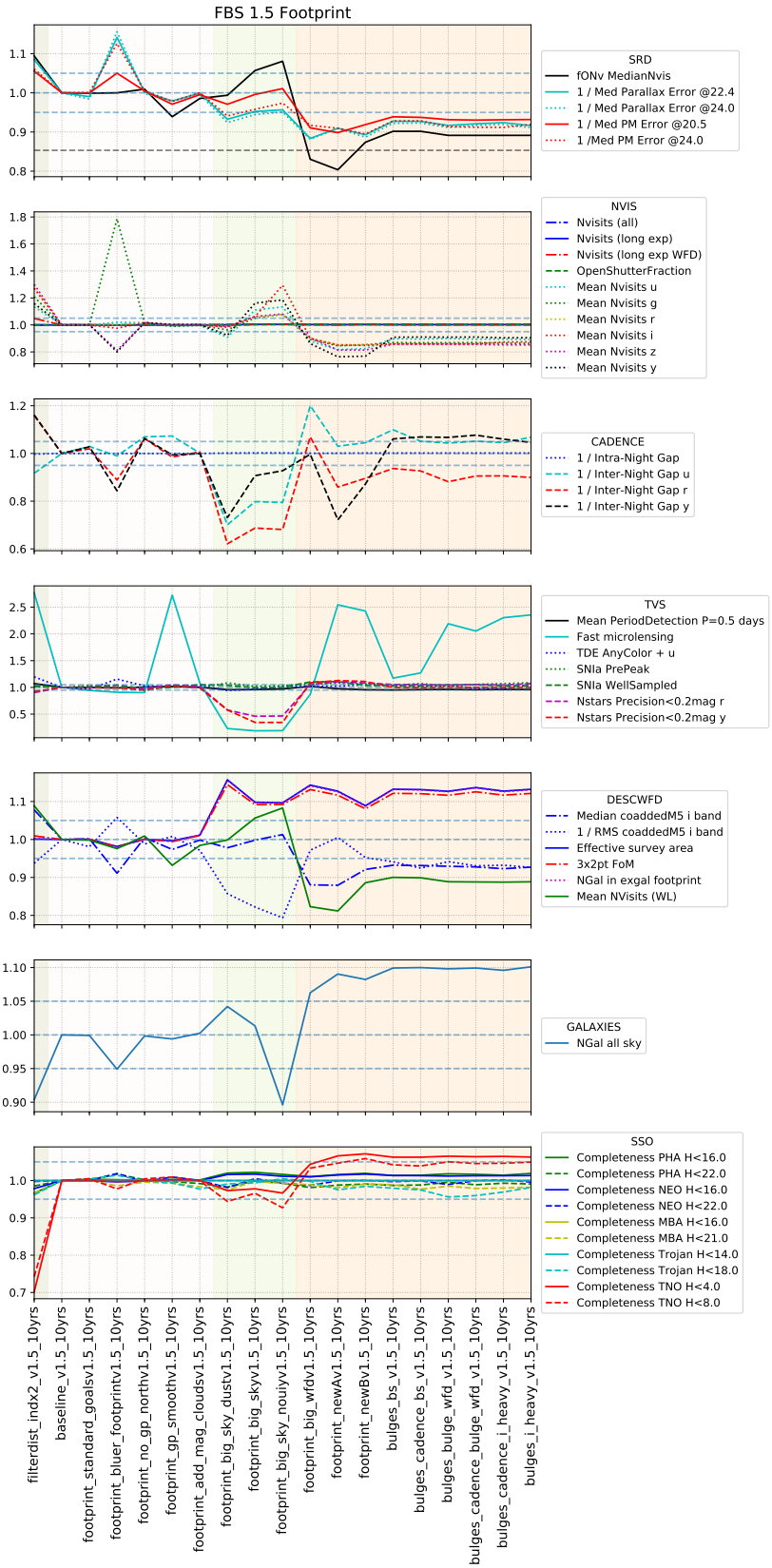

Figure 61. Comparing the simulations relevant to the footprint (Section 6.3). All metrics are scaled so that bigger is better and all are normalized against the baseline run. The dashed black line on the SRD metrics subplot indicates where fONv fails to meet requirements.

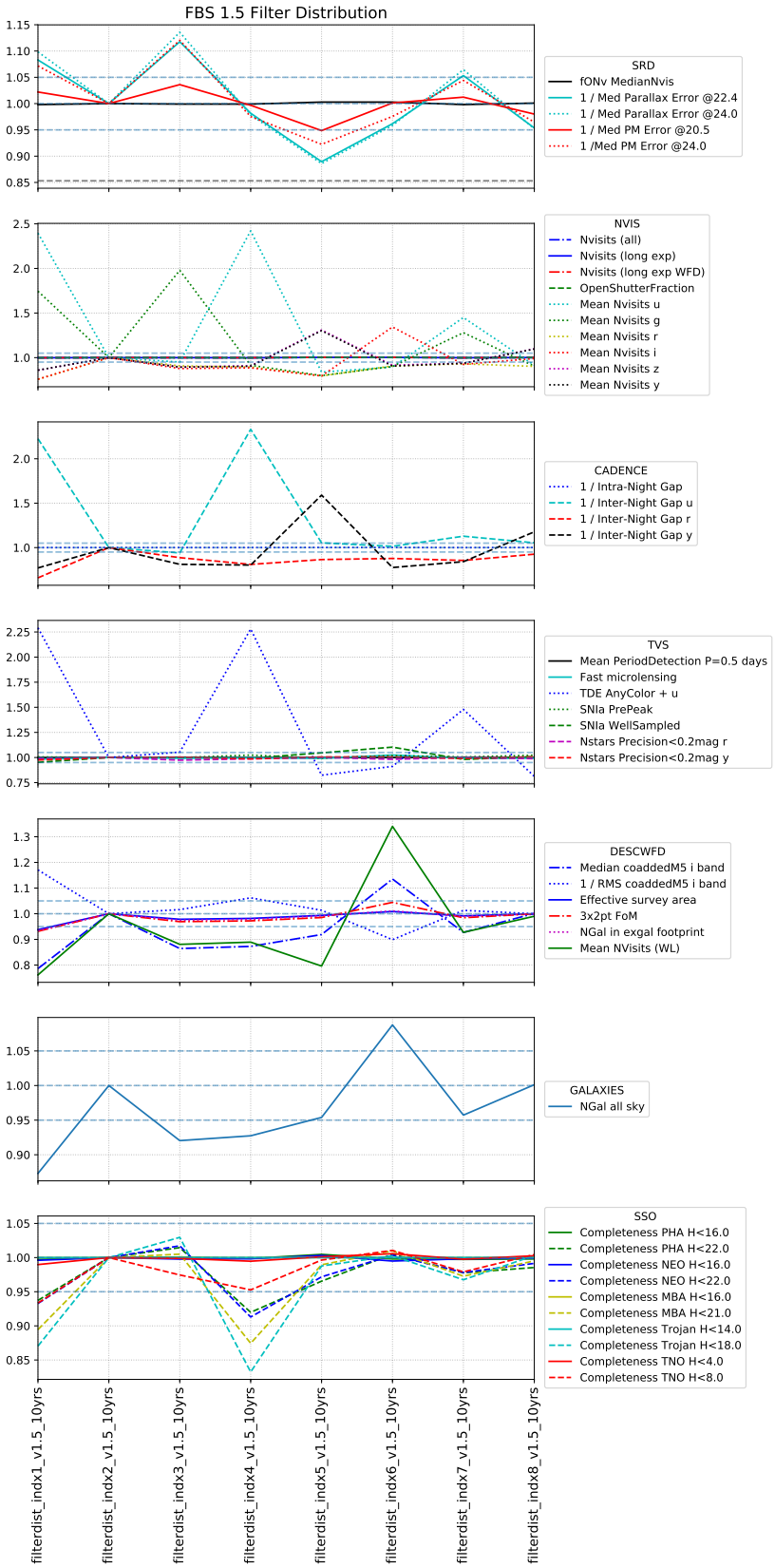

Figure 62. Comparing the simulations relevant to the filter distribution (Section 6.3). All metrics are scaled so that bigger is better and all are normalized against the `filterdist_indx2` run. The dashed black line on the SRD metrics subplot indicates where fONv fails to meet requirements.

The overall impact – whether additional minisurveys break SRD metrics or not, for example – will depend on what the general survey strategy is, as well. The background survey strategy for these experiments was a standard baseline, and we see from Figure 60 that the WFD in this kind of footprint generally runs above 80%; it has some 'room to spare'. A footprint similar to the big sky footprint, with NES, SCP, and GP coverage (such as the `newB` or `bulges` family) has a much lower fraction of visits in the WFD, about 70%. A minisurvey running with this background footprint would have a more immediate impact on SRD metrics. With the baseline survey stategy 'background', we see the `twilight_neo_mod1` minisurvey is the only case where the SRD metrics are broken – this minisurvey is removing about 7% of standard visits from the WFD, but has a strong effect on WFD-sensitive metrics, perhaps because it is also consistently redirecting the time distribution of WFD visits out of twilight and into the main portion of the night. The `short_exp_5ns_5` minisurvey drops WFD visits by about 6% but does not have the same impact, likely because those visits are slightly longer and at least to some level within the goals of visits in the WFD. Note how the minisurvey with 5s visits 2 times per year and the minisurvey with 2s visits 5 times per year have a similar impact on the number of standard WFD visits, but the 5s minisurvey has a smaller impact on metrics. The DCR minisurvey, which redirects standard WFD visits to high airmass visits, modifies even larger fractions of the WFD visits (up to15%), but has even less impact because these visits are still closer to the usual WFD goals.

To properly evaluate the DCR and short exposures minisurveys, we need additional metrics sensitive to these effects. For the twilight NEO survey, we should test an additional population of Vatiras. However, we can clearly see that minisurveys requiring more than about 2% of time (when these visits are not useful for the WFD, such as the `twilight_neo` family) have severe impacts on other science areas, with a standard baseline background. It is worth considering that the time used by these minisurveys could represent time lost to any cause and that this impact will be more immediate when running with a different general survey footprint.

6.5. *The FBS 1.6 runs*

The FBS 1.6 runs don't include (generally) include minisurveys, but are configured to spend more time overall on the WFD, even when running on footprints that are extensions of the 'big sky' WFD, such as the `combo_dust` simulations. Figure 65 shows that these runs range from spending 75% of time on the WFD up to 95% of time (in the `barebones` simulation). All of the resulting simulations meet SRD requirements when run with 1x30s visits; most of them also meet requirements when running with 2x15s visits, and those that fail are within a percent or so of the requirements. This suggests there is not a large amount of contingency for SRD requirements in these modes nor room to add minisurveys requiring beyond a percent or two of time; on the other hand, our weather model is reasonably conservative and there is room to

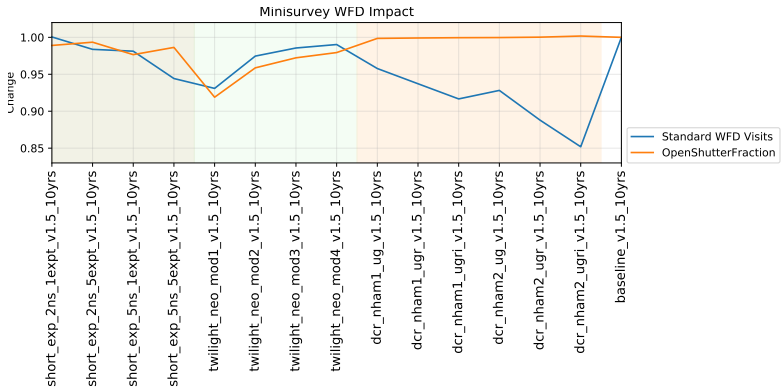

Figure 63. The amount of time required for each minisurvey is difficult to measure when they have variable exposure times. Because the remaining survey strategy parameters were constant, a useful analog is the change in the number of standard, non-minisurvey, visits for the WFD, although this is likely to be inexact. The fractional change in the number of WFD visits (compared to the baseline) when each minisurvey was was running is shown here. Some of the minisurveys require very little time; some require more and are visible in other science metrics as well; some require significant time but the type of visit overlaps (although is suboptimal for) the WFD.

modify the survey strategy if early years prove problematic.The simulations with 1x30s visits have accordingly larger margins with respect to SRD metrics. Metric results are shown in Figure 66. These runs show similar trends as we saw previously when considering the survey footprints, with the comforting addition that only the `barebones` run seems to really break the science metrics.

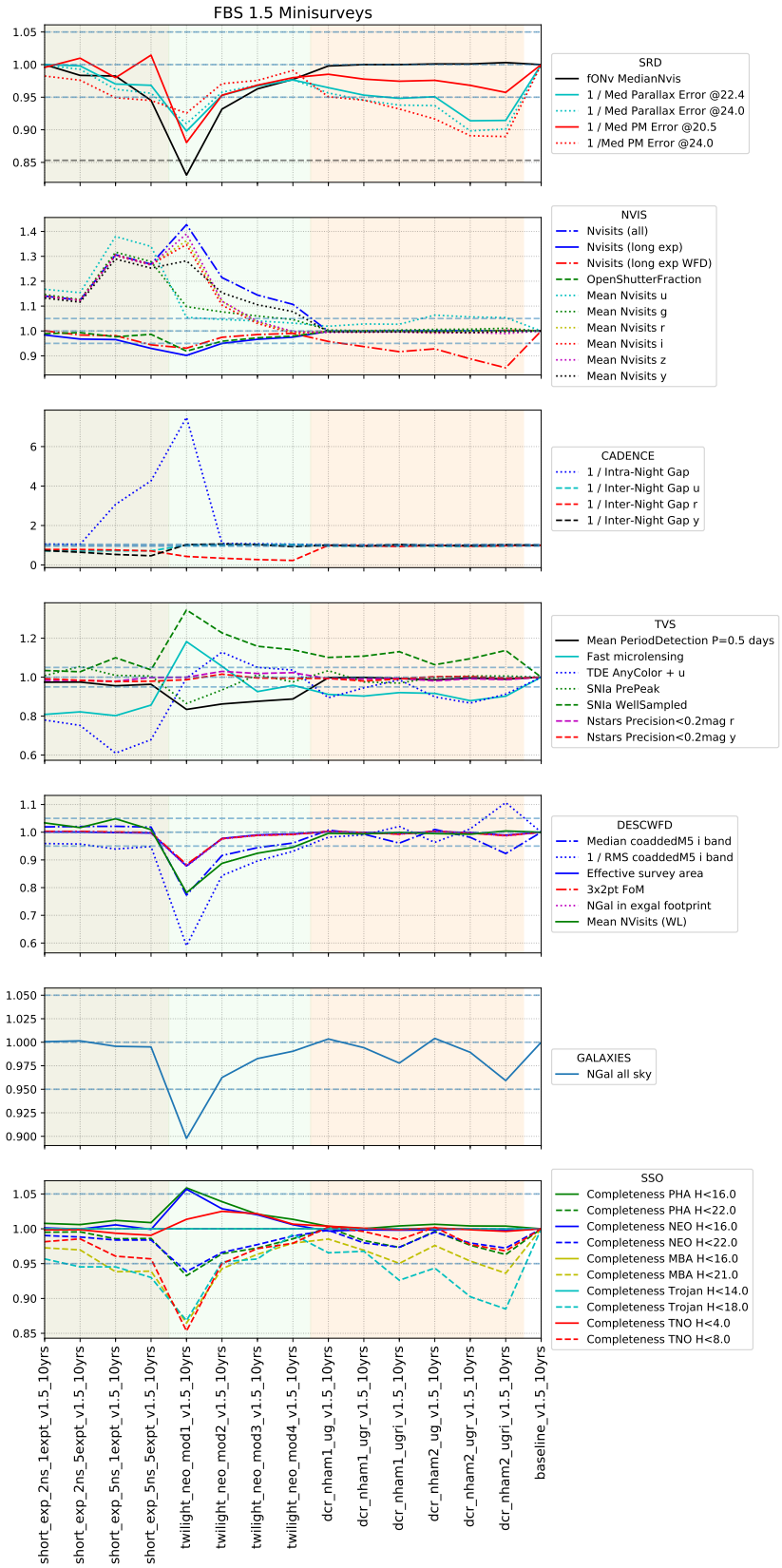

Figure 64. Comparing the simulations relevant to the addition of minisurveys (Section 6.4). All metrics are scaled so that bigger is better and all are normalized against the baseline run. The dashed black line on the SRD metrics subplot indicates where fONv fails to meet requirements.

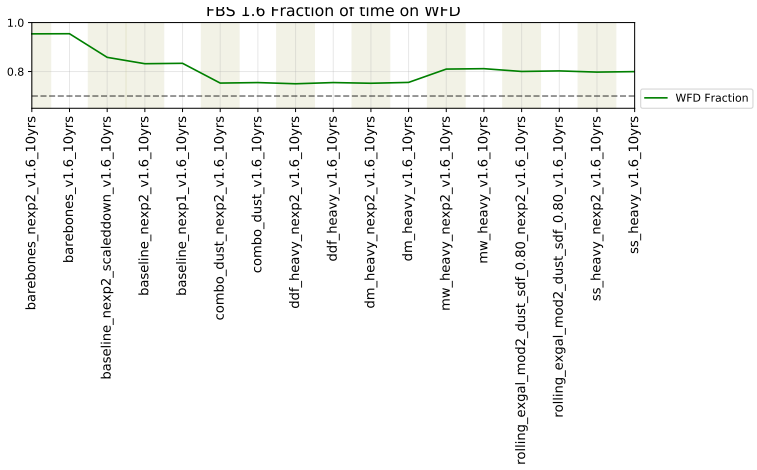

Figure 65. Comparing the fraction of visits used to observe the WFD and the total number of visits gives an appropriate fraction of time spent on the WFD. As the survey focus and footprint changes, this fraction changes significantly. Shaded regions indicate 2x15s simulations; unshaded regions indicate 1x30s versions of similar simulations (the 1x30s simulation is to the right of the 2x15s run).

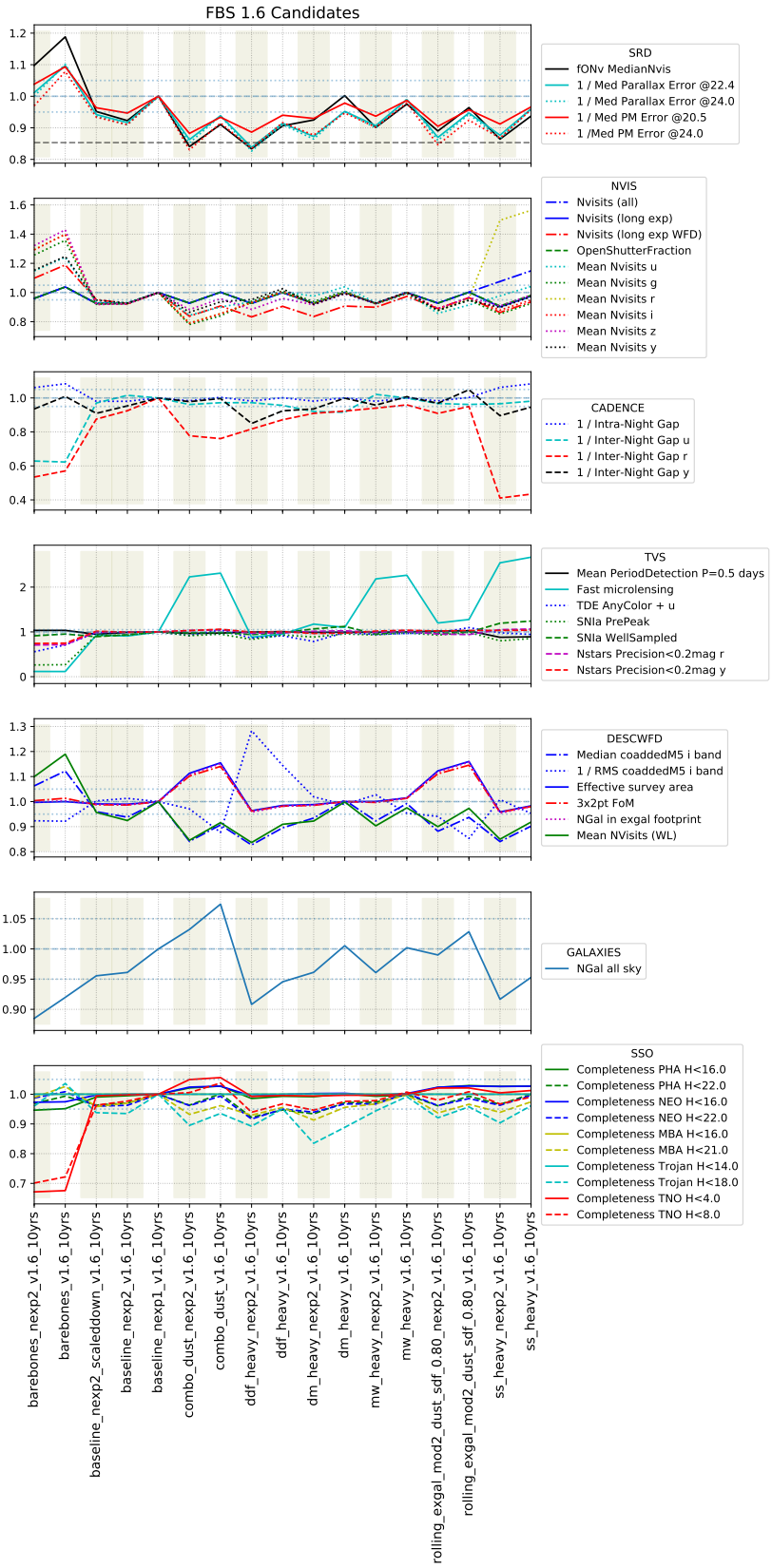

Figure 66. Comparing the v1.6 runs (Section 6.5). All metrics are scaled so that bigger is better and all are normalized against the baseline run. Shaded columns are 2x15s runs, unshaded are 1x30s runs. The dashed black line on the SRD metrics subplot indicates where fONv fails to meet requirements.

7. SUMMARY

Summing up the high-level questions addressed by these simulations –

Visit exposure time: We ran simulations with 2x15s visits and 1x30s visits, longer u band visits, and variable exposure time (Section 6.1). We cannot choose between 1x30s and 2x15s visits at this time. Reducing the number of u band visits in order to increase the exposure time is bad for transients; new simulations should be done once further survey strategy choices are made. Variable exposure times result in fewer visits over all, although similar on-sky exposure time; metric results were neutral.

Intra-Night cadence: We ran simulations with pairs in the same filter and mixed filters, as well as adding triplets for some portion of the night (Section 6.2. Mixed filters are preferable for most science and not significantly detrimental for any metrics. Adding triplets of visits for a small fraction of the sky (using 15 to 45 minutes at the end of each night) is relatively neutral for our metrics, perhaps due to pre-existing triplets in the pairs simulations due to field overlaps. More metrics sensitive to the presence or absence (and timing) of a third visit in a night would be useful.

Rolling cadence: We tested the rolling cadence using 2, 3 and 6 bands of declination-limited area (Section 4.9). There is an additional implementation of a 2-band rolling cadence that includes 4 N/S bands in Section 4.8, more fully explained in Section 5.5. Metrics are relatively neutral for the 2 and 3 band declination options. The 6 band option is mildly to strongly negative for Solar system science and some kinds of transient science (due to trades between the time and area coverage); it also is likely to be more vulnerable to weather disruptions. There are benefits to the alt-roll cadence that may outweigh any potential problems with the even-odd night cadence gaps, although aliasing may be a concern. More metrics sensitive to short-timescale transients and variables would be useful.

Filter distribution: We tested the effects of varying the filter distribution (Section 4.7 and Section 6.3). The SRD lays out a suggested filter distribution based on photometric redshift expectations; a photometric redshift metric to run on the simulations would be extremely useful. There is some tension between shifting to more bluer (u) visits vs. adding more redder (i) band visits.

DDFs: We ran simulations using various DD cadences (Section 4.10). Evaluating the effect of these simulations requires more specific DDF metrics. If the DDFs are capped at a total fraction of survey time, their impact on the remainder of the survey is consistent. If we allow their time requirement to vary depending on the cadence requirements, and attempt to meet both DESC and AGN DDF cadence requirements, more time will be required and their impact will be more significant.

Minisurveys: We tested various minisurvey additions – a DCR minisurvey (Section 4.16), a short exposure minisurvey (Section 4.13), and a twilight NEO minisurvey (Section 4.12), and evaluated the general impact of minisurveys in Section 6.4. Observations taken for minisurveys tend to be suboptimal for standard surveying, but some minisurveys have a smaller effect than others. Metrics which target the expected

science return of these minisurveys are needed to fully evaluate these trades. It is likely that decisions on the minisurveys may have to wait until after further choices about the survey footprint are made.

Survey footprint: The choice for the survey footprint has the greatest repercussions for the overall survey strategy (Section 6.3. The options can be split into two broad groups that are differentiated based on the WFD footprint – the current baseline WFD footprint which runs from $-62°$ to $2.5°$ declination with a relatively small cutout around the galactic plane and a 'big sky' style WFD footprint which runs from $-72°$ to $12.5°$ declination with a wider cutout around the galactic plane. Most metrics prefer variations on the big sky WFD where additional coverage is added for the NES, SCP and GP, but these runs result in fewer visits per pointing within the WFD region and have a lower margin to passing SRD metrics.

Over the next two years, the SCOC will look for further input from the community on metrics and on these survey strategy investigations. This process will necessarily be somewhat iterative; with additional feedback, an updated set of more targeted simulations can be created and evaluated using new metrics, which will feed back to the SCOC. Two community workshops are currently planned, as well as further communication directly between the SCOC and the science collaborations as well as the community at large. The public Survey Strategy forums at community.lsst.org will host additional thoughts and comments, and are a great place to ask questions or start discussions. We look forward to participating in this process to find the best survey strategy for the LSST.

REFERENCES

Awan, H., Gawiser, E., Kurczynski, P., et al. 2016, ApJ, 829, 50, doi: 10.3847/0004-637X/829/1/50

Bellm, E. C., Kulkarni, S. R., Barlow, T., et al. 2019, PASP, 131, 068003, doi: 10.1088/1538-3873/ab0c2a

Delgado, F., Saha, A., Chandrasekharan, S., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9150, Modeling, Systems Engineering, and Project Management for Astronomy VI, ed. G. Z. Angeli & P. Dierickx, 915015, doi: 10.1117/12.2056898

Girardi, L., Groenewegen, M. A. T., Hatziminaoglou, E., & da Costa, L. 2005, A&A, 436, 895, doi: 10.1051/0004-6361:20042352

Girardi, L., Barbieri, M., Groenewegen, M. A. T., et al. 2012, in Astrophysics and Space Science Proceedings, Vol. 26, Red Giants as Probes of the Structure and Evolution of the Milky Way, ed. A. Miglio, J. Montalbán, & A. Noels, 165, doi: 10.1007/978-3-642-18418-5_17

Granvik, M., Morbidelli, A., Jedicke, R., et al. 2018, Icarus, 312, 181, doi: 10.1016/j.icarus.2018.04.018

Grav, T., Jedicke, R., Denneau, L., et al. 2011, PASP, 123, 423, doi: 10.1086/659833

Ivezić, Ž., Jones, L., Ribeiro, T., LSST Project Scient Team, & LSST Science Advisory Committee. 2018, Call for White Papers on LSST Cadence Optimization, Informal Construction Document Document-28382, NSF-DOE Vera C. Rubin Observatory. https://ls.st/Document-28382

Ivezić, Ž., & The LSST Science Collaboration. 2018, LSST Science Requirements Document, Project Controlled Document LPM-17, NSF-DOE Vera C. Rubin Observatory. https://ls.st/LPM-17

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111, doi: 10.3847/1538-4357/ab042c

Jurić, M. 2018, galfast: Milky Way mock catalog generator„ Astrophysics Source Code Library, record ascl:1810.001

Jurić, M., Ivezić, Ž., Brooks, A., et al. 2008, ApJ, 673, 864, doi: 10.1086/523619

Kavelaars, J. J., Jones, R. L., Gladman, B. J., et al. 2009, AJ, 137, 4917, doi: 10.1088/0004-6256/137/6/4917

Lampoudi, S., Saunders, E., & Eastman, J. 2015, arXiv e-prints, arXiv:1503.07170, doi: 10.48550/arXiv.1503.07170

LSST Science Advisory Committee. 2019, Recommendations for Operations Simulator Experiments Based on Submitted Cadence Optimization White Papers, Informal Construction Document Document-32816, NSF-DOE Vera C. Rubin Observatory. https://ls.st/Document-32816

Naghib, E., Yoachim, P., Vanderbei, R. J., Connolly, A. J., & Jones, R. L. 2019, AJ, 157, 151, doi: 10.3847/1538-3881/aafece

Petit, J. M., Kavelaars, J. J., Gladman, B. J., et al. 2011, AJ, 142, 131, doi: 10.1088/0004-6256/142/4/131

Rothchild, D., Stubbs, C., & Yoachim, P. 2019, PASP, 131, 115002, doi: 10.1088/1538-3873/ab3300

Yoachim, P., Coughlin, M., Angeli, G. Z., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9910, Observatory Operations: Strategies, Processes, and Systems VI, ed. A. B. Peck, R. L. Seaman, & C. R. Benn, 99101A, doi: 10.1117/12.2232947

8. ACRONYMS