Rubin Cadence Note - Bulge stellar populations with LSST

Contributors | Motivation | Recommendations | SCOC Questions | Strategies considered | Figures of merit | Figures | References

Contributors

Will Clarkson (UM-Dearborn), Massimo Dall'Ora (INAF-OACN), Alessandro Mazzi (University of Padova), Leo Girardi (INAF), Victor P. Debattista (UCLan), Annalisa Calamida (STScI), Oscar Gonzalez (UKATC), R. Michael Rich (UCLA), Christian I. Johnson (STScI), Knut Olsen (NOIRLAB), G. Bono (University of Tor Vergata), G. Fiorentino (INAF-OAR), Xiaolong Li (University of Delaware), R. Szabo (Konkoly Obs.)

Acknowledgements: All calculations make heavy use of the **sims_maf** framework developed by Lynne Jones and Peter Yoachim. They were run in the NOIRLAB Datalab JupyterLab environment. We thank Peter Yoachim and Lynne Jones for their continuous, patient and comprehensive help and support with all aspects of the development of the metrics and figures of merit described here.

Motivation

The Galactic bulge is one of the most important topics in the Rubin-LSST key science area "Mapping the Milky Way," since several problems are still open about its 3D structure, its age-metallicity distribution and, in turn, its formation history. Here we discuss the impact of the different observing strategies on selected stellar tracers: a) RR Lyrae stars (RRL, for distances, reddening and metallicities); b) Red Clump stars (RC, distances and metallicities); c) Main Sequence Turn-Off stars (MSTO, for age distribution). Several papers (e.g. Pietrukowicz et al. 2015; Saha et al. 2019; Johnson et al. 2020) address the proper techniques to get the relevant information for each of the aforementioned tracers.

The leading challenges when observing the bulge with LSST are the strong, variable extinction, and the high degree of spatial crowding to seeing-limited imaging observations. There are thus two points to establish: (i) that LSST *can* deliver Galactic bulge stellar populations, and by extension crowded fields science at a useful level; and (ii) if so, which strategies are preferred. Point (i) is established by experience with previous facilities, which we use to calibrate the predictions of MAF for the various strategies to address point (ii). Several observational campaigns (e.g. with DECam on the 4m-Blanco telescope; Saha et al. 2019; Schlafly et al. 2018; Rich et al. 2020; Johnson et al. 2020) have demonstrated that a seeing-limited ground-based imager can make scientifically useful measurements of all three classes of tracer mentioned above by using a similar set of filters as LSST. Further, these papers used DECam on a 4m telescope, and the image quality and end of mission depth of LSST will be significantly better than for DEcam. This cadence note focuses on point (ii): which types of strategy are better for measuring bulge stellar tracers with LSST.

Although the Rubin Observatory is not ideally located to study the Northern bulge, we emphasize the unique value of Rubin data for that program. The Northern bulge is more affected by extinction, but the multicolor Rubin photometry would be ideal for extinction correction. Observations of the Northern bulge would allow studies of the structure of the bulge as a function of metallicity and tests of whether unusual structures such as the X-shape, changes in the metallicity distribution etc. are found in the Northern bulge.

Recommendations

1. Deep u-band imaging is crucial for bulge science; more time in u-band is generally better. For the same total exposure time, 1x50s in u is preferable to 2x15s.
2. Strategies that extend Wide-Fast-Deep (WFD) to the plane are preferred, with uniform coverage of the inner plane preferred to strategies that are “dust-limited.”
3. For tracing populations through variability, filter uniformity (among [g,r,i,z,y]) is preferred over cadences that concentrate observations in one or two filters.
4. Strategies allocating 180 or more exposures in total to the bulge are preferred, with at least some coverage at the beginning and end of the 10-year survey to afford sensitivity to long-term variability and proper motion.

We would also request that opsims be run that allocate good seeing observations towards crowded regions, including the bulge and Magellanic Clouds. At this date we do not yet have detailed specifications for these experiments, but we expect to develop them in the near future.

Responses to SCOC queries

Q1: Are there any science drivers that would strongly argue for, or against, increasing the WFD footprint from 18,000 sq. deg. to 20,000 sq.deg.? Note that the resulting number of visits per pointing would drop by about 10%. If available, please mention specific simulated cadences, and specific metrics, that support your answer.

Extension of WFD to include the bulge is preferred: within v1.7.1, the **full_disk_*** opsims perform the best for population recovery via integrated depth, colors, and in supporting variability (Figure 1). Extension of WFD in a way that does *not* include the bulge performs slightly less well for these figures of merit (**footprint_big_wfdv1.5_10yrs**).

Q2: Assuming that current system performance estimates will hold up, we plan to utilize the additional observing time (which may be as much as 10% of the survey observing time) for visits for the mini-surveys and the DDFs (with an implicit assumption that the main WFD survey meeting SRD requirements will always be the first priority). What is the best scientific use of this time? If available, please mention specific simulated cadences, and specific metrics, that support your answer.

The bulge minisurvey is fundamental for our science. The Galactic bulge is a key actor in the formation history of the Milky Way galaxy, and represents an important unsolved area of one of the four key science drivers for the Rubin Observatory. The structure of the bulge as a function of metallicity and age, and its relationship to the inner disk is far from well defined. A uniform photometric catalog of the bulge using LSST’s capabilities, will produce many independent probes of the bulge’s present day disposition and formation history, for a relatively modest allocation of a few percent of the total survey time. With a stellar mass of $2 \times 10^{10}\ M_{sun}$ and high stellar density, the Galactic bulge is also likely to be the host of numerous and interesting transient events, such as classical novae, luminous red novae, and possible new classes of transient events (explored in more detail in cadence notes to be submitted by TVS, e.g. Bachelet et al.). The high image quality and regular cadence operations of LSST will provide data of unmatched quality for these investigations.

Q3: Are there any science drivers that would strongly argue for, or against, the proposal to change the u band exposure from 2x15 sec to 1x50 sec? If available, please mention specific simulated cadences, and specific metrics, that support your answer.

[Fe/H] measurements are crucial to disentangling the formation history of the inner Milky Way, and require u-band coverage. Our volume-to-(u-i) figure of merit demonstrates that longer single u-band exposures are preferred for photometric metallicity estimates; for example, opsims **u_long_ms_50_v1.710yrs** outperforms the relevant baseline cadence for [Fe/H] estimates.

Q4: Are there any science drivers that would strongly argue for, or against, further changes in observing time allocation per band (e.g., skewed much more towards the blue or the red side of the spectrum)? If available, please mention specific simulated cadences, and specific metrics, that support your answer.

According to the RRL metrics, any opsim skewed toward a particular filter (i.e. u_long, but also i_heavy, significantly affects the RRL performance. This is because they need as much uniform coverage as possible in at least 4 bands. More generally, it appears that any technique relying on the combination of different bands is sensible to a uniform coverage in the employed filters. The exception is the determination of [Fe/H] via (u-i) color: for that science case, u-band depth is essential.

Q5: Are there any science drivers that would strongly argue for, or against, obtaining two visits in a pair in the same (or different) filter? Or the benefits or drawbacks of dedicating a portion of each night to obtaining a third (triplet) visit? If available, please mention specific simulated cadences, and specific metrics, that support your answer.

Measuring depths, colors, and RR Lyrae variability, do not strongly argue for pairs or triplets, as long as sufficient total observations are obtained, and as long as the full 10 years of the survey are covered. For the tracer populations considered here, colors will be estimated from the entire lightcurves in the different filters (for variables like RR Lyrae as well as for non-variable objects).

Q6: Are there any science drivers that would strongly argue for, or against, the rolling cadence scenario? Or for or against varying the season length? Or for or against the AltSched N/S nightly pattern of visits? If available, please mention specific simulated cadences, and specific metrics, that support your answer.

Rolling cadences do not perform worse than baseline from the perspective of bulge tracers (either static or for variability as traced by RR Lyrae and similar variables), although rolling cadences are generally **not** the best-performing strategies from this perspective (e.g. Figure 1). More impulsive or short-term variability (like transients or short-timescale microlensing) would likely benefit from rolling cadence, but other cadence notes cover this ground in more detail. An extended season length has the potential to detect a greater number of transients.

Of the v1.5 strategies we tested, the AltSched strategy performs the most poorly in terms of the figures of merit considered here. (The restricted spatial region over which we evaluate the figure of merit is the same for all strategies, so this strategy should be directly comparable to the other strategies within v1.5.)

Q7: Are there any science drivers pushing for or against particular dithering patterns (either rotational dithers or translational dithers?) If available, please mention specific simulated cadences, and specific metrics, that support your answer.

N/A

Strategies considered

At the time of writing, 34 candidate strategies have been examined for bulge stellar populations, covering a range of choices of footprint, cadence, and distribution among filters:

Program exposure length	**baseline_nexp1_*** **baseline_v1.5_10yrs**	(for v1.7, 1.7.1): 2x15s in each filter replaced with 1x30s (For v1.5, the baseline exposure length is 1x30s)
Exposure time in u	**u_long_ms_50_v1.7_10yrs**	2x15s u-band exposures replaced by 1x50s
WFD in the plane	**footprint_8_v1.710yrs** **bulge_roll_scale0.90_nslice***	Wide-fast-deep (WFD) across the bulge, “dust-limited” As above with rolling cadence implemented
	footprint_6_v1.710yrs **footprint_6_v1.7.1_10yrs**	WFD across the bulge, “dust-limited” + high-longitude bridge As above with rolling cadence implemented
	full_disk_* **footprint_gp_smoothv1.5**	WFD for -60 < dec < 0
Total survey area	**footprint_big_wfdv1.5_10yrs**	WFD extended to high declination; inner plane coverage unchanged from baseline
Rolling cadence	**rolling_nm_*** **six_stripe_***	As baseline but including rolling cadence in various ways
Big sky + bulge	**bulges_***	Combinations of bulge coverage + WFD or “big sky”
Filter balance	**bulges_*_i_heavy***	More i-band observations towards the bulge
	filterdist_indx*_v1.5_10yrs	Filter balance varied between the filters [indx1, 2 are uniform and baseline, indx 3, 4, 5, 6 heavy in g, u, z&y, i respectively. Indx7, indx8 are “bluer” and “redder”.
Seeing	**goodseeing_gi_v1.5_10yrs**	Prioritize achieving at least one “good seeing” image in g,i per year during the 10 years
AltSched	**alt_dust_1.5_10yrs**	Opsim implementation of the AltSched algorithm

Figures of Merit

Assessment of bulge observation strategies requires accounting for spatial confusion and for the strong, variable extinction over the region of interest. After clarifying the appropriate parameters, we have verified that the MAF predictions for the LSST spatial confusion limits based on the TRILEGAL simulations, do roughly match those achieved by DECam towards the bulge, in the areas in which previous data are confusion limited; this comparison will be communicated elsewhere: based on this experience, we expect that the MAF predictions for spatial confusion should be reliable to about 0.5 magnitudes, and we employ MAF’s confusion limit predictions throughout. To convert apparent magnitude limits into distance limits for assumed absolute magnitude, we employ a mixture of Bovy’s **mwdust** (Bovy et al. 2016), the new Gaia DR2-based dust-maps of Lallement et al. (2019), and (for any sight-lines for which **mwdust** predicts zero extinction) the 2D extinction map from the Planck collaboration.

All figures of merit in this communication are evaluated over the spatial region encompassing Galactic longitude | l | < 20^0 and latitude |b| < 9^0, show in Figures 1b & 2b.

Applying metrics that include 3D dust and spatial confusion (in all filters), we implement the figures of merit listed below.

1. [Fe/H] of Red Clump Giants

Figure 1 shows the Red Clump Giant [Fe/H] figure of merit. The metric evaluates the distance at which an RCG (absolute magnitude M_r ~ +0.55; e.g. Ruiz-Dern et al. 2018) would be measured with uncertainty in (u-i) of 0.18 dex or better. By expression (22) of Johnson et al. (2020), and assuming for the sake of strategy comparison that the extinction is perfectly known, this uncertainty threshold roughly corresponds to a 0.1 dex uncertainty in photometric [Fe/H] requiring that LSST photometry adds less to the error budget than the ~0.2 dex intrinsic scatter in the [Fe/H] - $(u-i)_0$ relationship of Johnson et al. (2020). While the dependence on observing strategy (among those tested) appears to be at the 25% level, this does suggest that [Fe/H] measurements of RCGs by LSST will be possible for most fields for |b| < ~ 7^0.

This estimate may be somewhat conservative: (i) Johnson et al. (2020) demonstrated the [Fe/H] vs $(u-i)_0$ relationship for a field at (l,b) = (+1,-4) from DECam data; and (ii) for fields in which observations in both filters are confusion-limited, the uncertainty on the color is *lower* than that predicted by treating the filters separately (Olsen, Blum & Rigaut 2003). At the time of writing, a confusion limit for color has recently been implemented (by KO); we plan to update this figure of merit using this new capability in the near future.

Of the opsims considered here, the (u-i) figure of merit varies by about 25%. As might be expected, strategies do better by this figure of merit that do any of the following:

- Spend more time in u-band (e.g. **filterdist_indx4_v1.5**) or use longer u-band exposures (e,g, **u_long_ms_50_v1.7_10yrs**); and/or
- Allocate WFD-type coverage towards the bulge, without extinction gaps (e.g. **full_disk_***).

Strategies that allocate increased i-band coverage towards the bulge (the ***_i_heavy_*** opsims) do not outperform the baseline strategy (for v1.5) from the perspective of (u-i) color. The **goodseeing_gi** opsim also does not outperform the baseline (for v1.5) despite the influence of spatial confusion in these regions. This is probably because the confusion metric here has been estimated using the best seeing from the set of exposures in any particular sight-line (and which probably therefore includes an extra contribution due to the total number of exposures: a larger number of exposures is more likely to sample an image or two of good seeing).

It would be useful to explore the impact (on bulge science as well as on other programs) of strategies that allocate best-seeing exposures to particular spatial regions - perhaps those in which the spatial confusion is predicted to be greatest. At this date we do not have a detailed recommendation for these simulations, but we hope to specify them in the near future.

2. RR Lyrae & multi-color precision

As shown in Bono et al. (2019) and proposed in the Bono et al. 2018 white paper (link), multiband photometry can be used to simultaneously measure individual distances, reddening and metallicities of RRLs (dubbed the *Redime* method). While the observing cadence will not have a major impact on the detection and on the characterization of the *individual* RRLs (see the Cadence Note - CN - *Classical variable stars in different Galactic environments: pulsation behaviour recovery*, by Musella et al.), nevertheless it will impact the number of RRLs measured with the needed photometric accuracy to properly use this new methodological approach, and therefore the bulge volume that can be probed.

As extensively discussed in Bono et al. (2019), the *Redime* method takes advantage of moderately high-precision optical-NIR photometry of RRLs, of the order of 0.04 mag in all the employed bands. If we assume that the photometric accuracy is the difference between a *true* RRL mean magnitude and the *observed* one from its lightcurve, then this accuracy is easily achievable even after 4-6 years of survey (see Musella et al. CN, their figure 8, where the comparison is made on simulated light curves). For our experiment, we ran two different simulations, one with the *grizy* set of filters, the other with only the reddest *rizy* bands, in order to avoid the stronger absorption affecting the *g*-band, which in turn limitates the Bulge volume at reach.

Figure 2 shows the volume that can be probed to RR Lyrae, at the declared photometric accuracy in the five *grizy* and four *rizy* bands.

On the basis of the opsims considered here, the probed volume varies by about 20%. Strategies with better performances are those with a quite “flat” coverage of the whole sky, with slightly different flavors, such as the **filterdist_indx5_v1.5_10yrs** and the **fulldisk_v1.7_10yrs**. This is not surprising, because the method is focused on the accuracy of the photometry, rather than on a large number of epochs (i.e. single visits). Interestingly, present simulations do disfavor *ad-hoc* **bulges*** strategies. Finally, we do not find any significant difference between the five *grizy*- and the four *rizy*-bands strategies. This means that the use of the *g* band is not a limiting factor for the method, because of the reddening limit on the latitude, which affects both *g*- and *r*-bands. (This is not to say that g-band photometry is not required; just that the amount of time allocated to g-band does not strongly impact the RR Lyrae figure of merit.)

3. Sufficient sampling for variability

We employ a simple metric to determine whether a field is sufficiently well-sampled that variability on RR Lyrae timescales can be constrained. Based on prior experience (e.g. reference XYZ and the Musella et al. cadence submission), a total of 80 exposures (or an average of just over 13 per filter over ten years) would allow “acceptable” period recovery, and 180 exposures (30 per filter over ten years) would allow “good” recovery (see for example Clementini et al. 2019; Stinger et al. 2019). Figures of merit for variability recovery to “acceptable” and “good” recovery fidelity, are estimated by summing the TRILEGAL star density predictions (to r = 22 mag) for all fields that pass the “acceptable” or “good” sampling criteria, over the region of interest corresponding to the bulge. All the opsims tested here perform essentially identically under this figure of merit (variation is less than 1%), and we conclude that all the strategies tested here are likely to allow sufficient sampling to constrain variability of tracer populations towards the bulge.

For a more continuous figure of merit, we quantify the total survey grasp to variability within the bulge region of interest, as (number of exposures x number of stars in TRILEGAL at r <= 22); larger values indicate better performance for variability. (The same bulge region is used for all strategies here, so the footprints of evaluation are identical.) This figure of merit allows us to identify strategies that cover more stars in the bulge region with sensitivity to variability. Figure 3 shows the results; since all the strategies considered here do allocate a sufficient minimum of exposures for variability, this figure of merit prefers strategies with larger allocations to the regions of interest.

(A more sophisticated figure of merit would consider the signal to noise per exposure for the typical tracer object, as well as the epoch distribution, but we have not implemented this at present.)

4. MSTO coverage and proper motion precision

The figure of merit described in Gonzalez et al. (2018) has been implemented in MAF, subject to a few updates. Based on experience comparing the observed BDBS luminosity function (Rich et al. 2020, Johnson et al. 2020), the maximum allowable crowding uncertainty has been updated to 0.25 to match the typical turnover in DECam data in crowded regions. The proper motion precision has been relaxed slightly, to 1 mas/yr; with large samples this is still useful for kinematic separation in crowded regions. At the present date this is still running; if indications are not complete by the Cadence Note deadline, we will update our indications as soon as possible after the deadline.

Figures

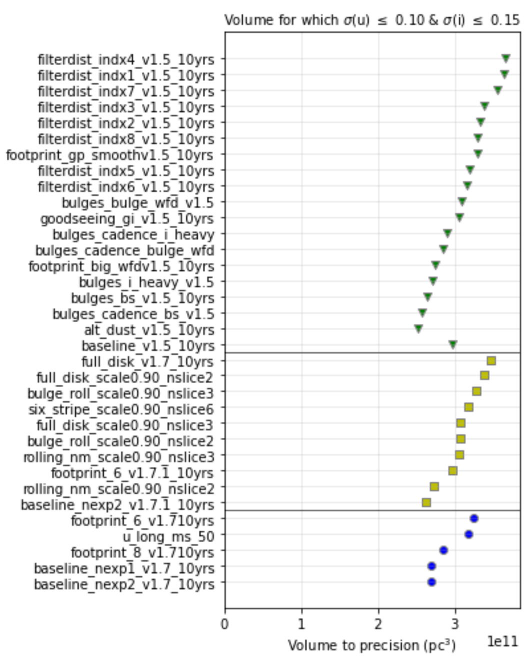

Figure 1a: Distance (in parsecs) at which a Red Clump Giant would be observed by LSST with (u-i) photometric uncertainty of 0.18 mag or better, which is our adopted threshold for [Fe/H] measurement from photometry via the calibration of Johnson et al. (2020). This figure of merit is expressed as a volume (towards the bulge; see Figure 1b) within which photometric estimates of [Fe/H] for RCGs are accessible to LSST. Larger numbers are better.

Opsims are partitioned by release, with the baseline opsim for that release appearing first, and the opsims are sorted within each release in order of increasing figure of merit.

Reading bottom-top, releases are: FBS 1.7, FBS 1.7.1, FBS 1.5.

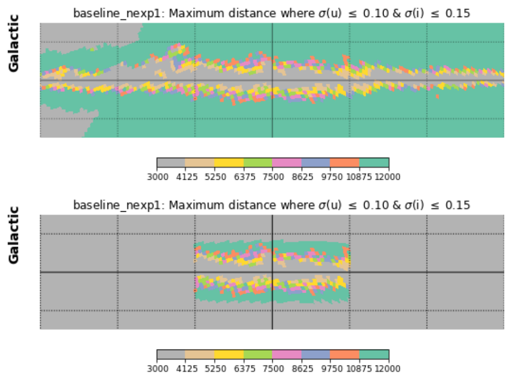

Figure 1b: sky-map for the baseline cadence (FBS 1.7), showing the region | l | < 60 and |b| < 15 degrees. The right-lower panel indicates the selection region for the bulge for which the figure of merit was computed (i.e. for |l| < 20^0 and between curved limits in |b|). In these figures, vertical dividers (meridians) are spaced 20 degrees apart, while horizontal dividers (parallels) are spaced every ten degrees.

Distance color shading is pegged to the (3-12 kpc) range. Reading left-right, the colorbar gradations take place at {3000, 4125, 5250, 7500, 8625, 9750, 10875 and 12000} pc, respectively.

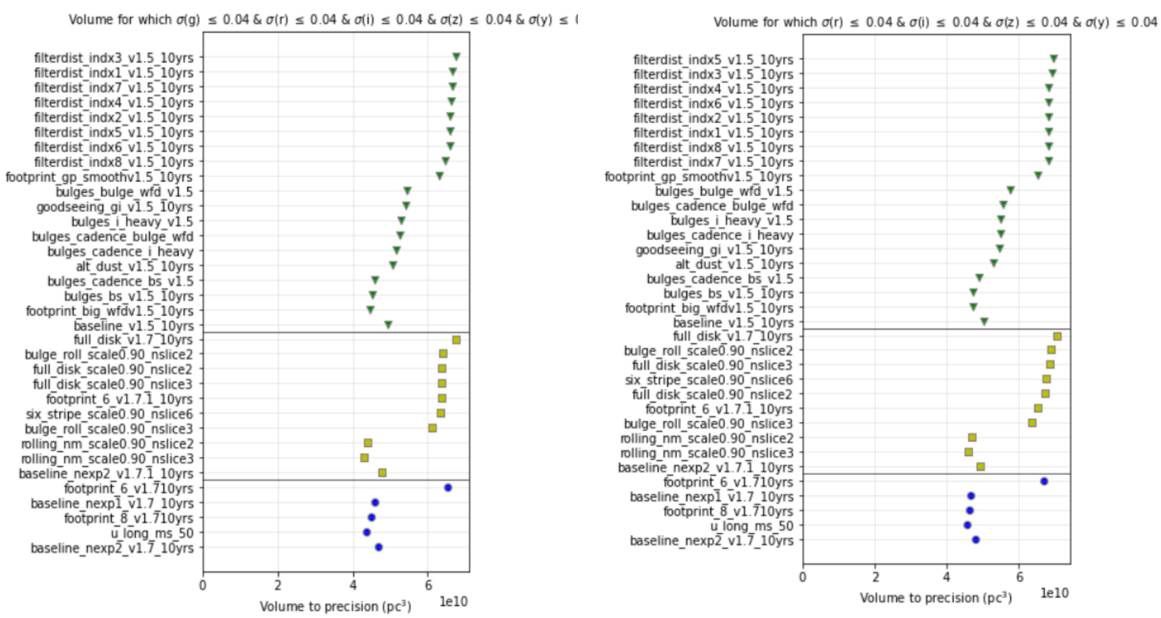

Figure 2a: Figure of merit constructed from the distance (in parsecs) at which a RR Lyrae star would be observed by LSST with photometric uncertainty of 0.04 mag or better in the *grizy* (left panel) and *rizy* (right panel), which is the recommended threshold to measure RRLs individual distances, reddening and [Fe/H] with an accuracy of 0.2 dex, from Bono et al. (2019). This figure of merit is expressed as a volume (towards the bulge; see Figure 2b) within which photometric estimates of [Fe/H] for RRLs are accessible to LSST. Larger numbers are better.

Opsims are partitioned by release, with the baseline opsim for that release appearing first, and the opsims are sorted within each release in order of increasing figure of merit.

Reading bottom-top, releases are: FBS 1.7, FBS 1.7.1, FBS 1.5.

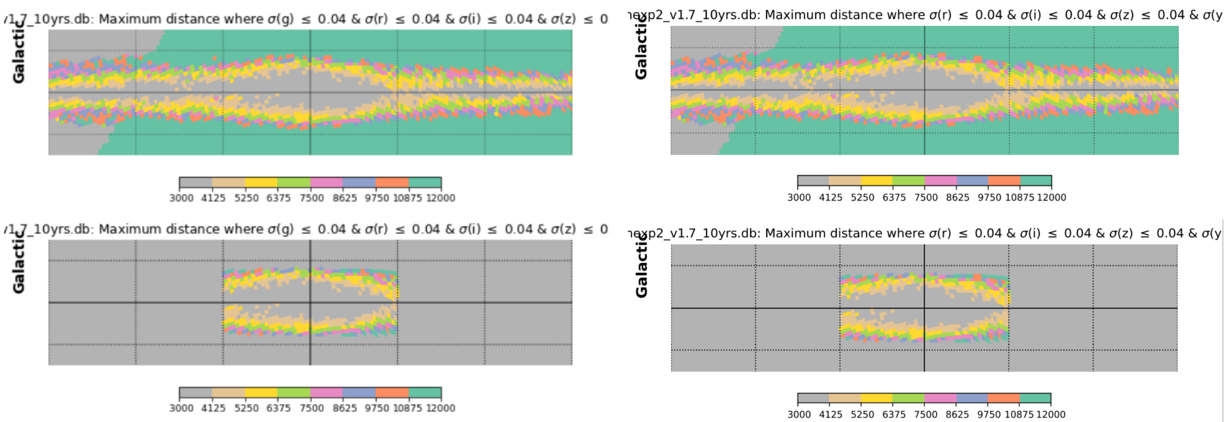

Figure 2b: sky-maps for the baseline cadence (FBS 1.7), showing the region | l | < 60 and |b| < 15 degrees. The lower panels indicate the selection region for the bulge for which the figure of merit was computed (i.e. for |l| < 20^0 and between curved limits in |b|). Distance color shading is pegged to the (3-12 kpc) range. In these figures, vertical dividers (meridians) are spaced 20 degrees apart, while horizontal dividers (parallels) are spaced every ten degrees.

Left panels show the results for the five *grizy* bands, while right panels are for the four *rizy* bands.

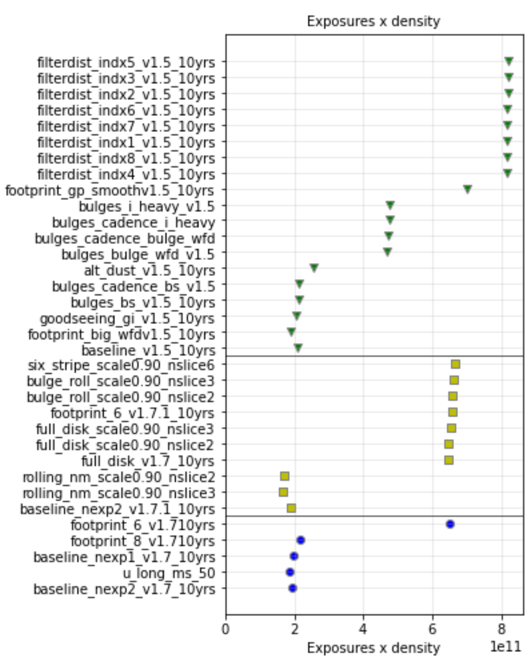

Figure 3: a simple figure of merit for survey grasp to variability within the bulge region of interest on the sky. Strategies are partitioned in a similar scheme to Figure 1; the baseline for each of the opsim versions appears at the bottom of each region, and strategies are ordered in increasing order of this figure of merit. Larger numbers are better.

References

Bono, G. et al. 2018 Cadence white paper

Bono, G. et al. 2019 ApJ, 870, 115

Bovy, J. et al. 2016 ApJ 818, 130

Clementini, G. et al. 2019 A&A, 622, 60

Gonzalez, O. et al. 2018 cadence white paper

Johnson, C. I. et al. 2020 MNRAS 499, 2357

Lallement, R. et al. 2019 A&A 625, A135

Lim, D. et al 2021 ApJ 907, 47

Olsen, K, Blum, R. & Rigaut, F., 2003 AJ 126, 452

Pietrukowicz, P. et al 2015 ApJ 811, 113

Rich, R. M. et al. 2020 MNRAS 499, 2340

Ruiz-Dern, L. et al. 2018 A&A 609, A116

Saha, A. et al. 2019 ApJ 874, 30

Schlafly, E. et al. 2018 ApJS 234, 19

Stringer et al. 2019 AJ 158 16

Wegg, C & Gerhard, O 2013 MNRAS 435, 1874