LSST AGN SC Cadence Note: Differential Chromatic Refraction

Weixiang Yu and Gordon Richards – on behalf of the AGN SC

1. EXECUTIVE SUMMARY

We have developed a metric to evaluate both the relative and absolute signal that can be expected from differential chromatic refraction (DCR) in LSST (Yu et al. 2020a) as a function of OpSims from versions FBS 1.5, 1.6 and 1.7. While there are LSST survey simulations that would significantly enhance this metric to the benefit of AGN classification and photo-z (and the study of objects exhibiting strong emission lines), such OpSims were meant as tests and not for operations. Among the feasible choices for LSST operations, we identify no simulations that are particularly bad or particularly good when it comes to DCR. Thus the choice of cadence can largely be made without consideration of this work. Nevertheless the tools are in place to double check that this remains true for the finalist(s).

2. THE METRIC

Fundamentally DCR is the positional offset (relative to the zenith direction) of the observed position relative to the true position. To first order this positional offset, Δ, is linear in the tangent of the zenith angle, Z: $\Delta = m\tan(Z) + b$. The intercept, $b \equiv 0$, as the DCR effect goes away at the zenith, $Z \equiv 0$. The slope, m, is set by the effective wavelength of the bandpass (which depends on the source SED). The accuracy to which we can measure the slope is largely *independent* of the source SED, with dependence on the astrometric error of the survey and the airmass of the observation. Both of these quantities are available at every location on the sky in the LSST OpSims (where the astrometric error is a function of magnitude and atmospheric seeing). Thus the error on the slope is (to first order): $\sigma_m = \sigma_{\rm astrom}/\tan(Z)$.

With this realization, we can make differential comparisons of LSST OpSims. The code is available for inspection from Yu et al. (2020b) and more details are available in the Yu et al. (2020a) RNAAS article. Note that our analysis is non-parametric and independent of the spectral energy distribution (SED) of the source. That is, we do not need to know *how* DCR information is going to be reported in order to know what information is potentially *available* or to know the SED of the source to know the *relative* effect that DCR will have between different simulations.

In the left panel of Figure 1 we present the results (in relative terms) for one member of each OpSim "family" (with the best metric for that family) and 5 baseline OpSims (that are used as comparison runs) over the Wide-Fast-Deep (WFD) survey. In the right panel we instead show DCR metric distributions for the 10 best and 10 worst OpSims (regardless of family). In Figure 2 we plot the median metric at each Deep-Drilling Field (DDF) with the metric values normalized to the result obtained at the "XMM-LSS" field in the `baseline_v1.5_10yrs` run. Every OpSim was summarized by the sum of the median metric at each DDF. The left panel of Figure 2 displays the 10 best and 10 worst OpSims in addition to all the OpSims that are designed specifically to test DDF cadence strategies (excluding those in the `euclid_dithers` family) at $u = 22.15$. The right panel of Figure 2 does the same for $g = 22$.

3. ANALYSIS OF RESULTS

The 10 OpSims that scored the highest in WFD include 6 versions from the `dcr` family, which "intentionally schedules a subset of images at high airmass" (PSTN-051)—as might be expected. Unsurprisingly the OpSims that have 2 high airmass observations are better than those with 1 and more filters receiving high airmass gives a higher DCR metric (even including r and i although the improvement is not significant). We find that the DCR metric is highest for two of the `filter_dist` experiments (`indx1`='uniform' and `indx3`='g heavy'). This family is "a simple WFD-only" simulation designed for testing photometric redshifts by varying the weights between the different filters. The `footprint_bluer` OpSim also scored well due to its "bluer filter distribution". The `dm_heavy` (v1.6) OpSim scored well due to extra high-airmass observations in ugr as did `barebones` (v1.6), although it is not considered a viable survey strategy.

The worst DCR metrics in the WFD were seen in three versions of the `twi_neo` and two versions of the `twi_pairs` simulations from v1.7, which attempt to search for NEOs and pair twilight observations, respectively, using 2×15 sec exposures. These OpSims seem not to meet the SRD requirements. Also scoring poorly were three `cadence_drive` simulations, which try to avoid long gaps in the g-band, perhaps resulting in fewer high-airmass observations. Fewer high-airmass observations are also likely the cause of a low metric in the `pair_times_11_v1.7` simulation, while the poor performance of the `ss_heavy` simulation might be expected.

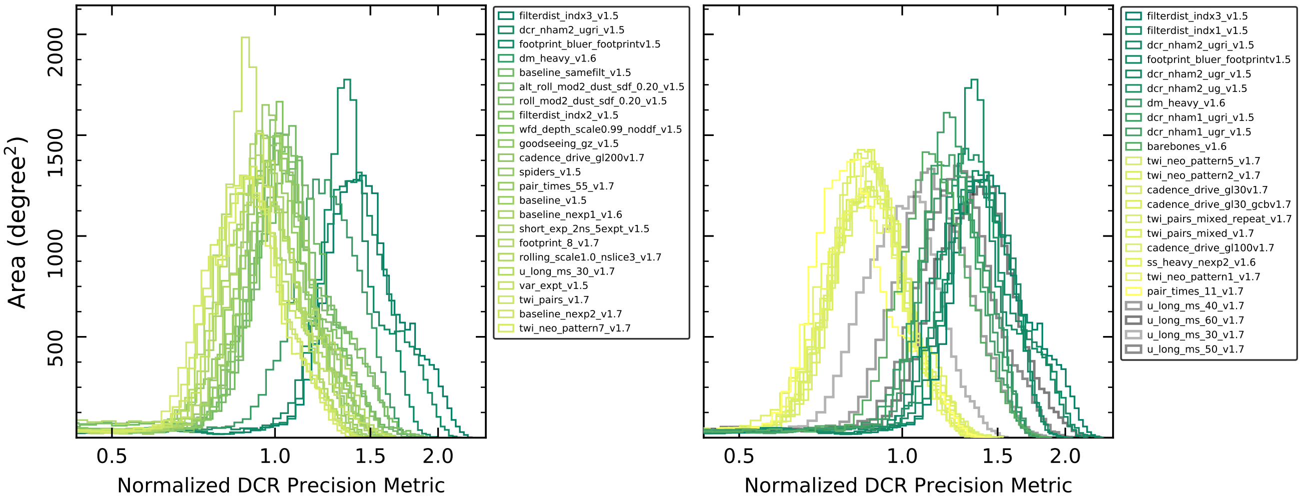

Figure 1. *Left:* Distribution of relative improvement in DCR slope uncertainties (on a log axis and normalized to the median metric of the `baseline_v1.5_10yrs` run) for 23 different LSST OpSims, using a fiducial depth of $g = 22$. The 23 different OpSims include one from each family (with the highest median metric) that vary the WFD survey strategy and the five baseline comparison runs. These results are independent of source SED. *Right:* The same distributions are plotted for the best 10 and worst 10 OpSims of all relevant OpSim runs. Also shown (in gray) are the distributions of the DCR metric evaluated at $u = 22.15$ for the OpSims in the `u_long` family. The same normalization run, `baseline_v1.5_10yrs`, was used as in the g band.

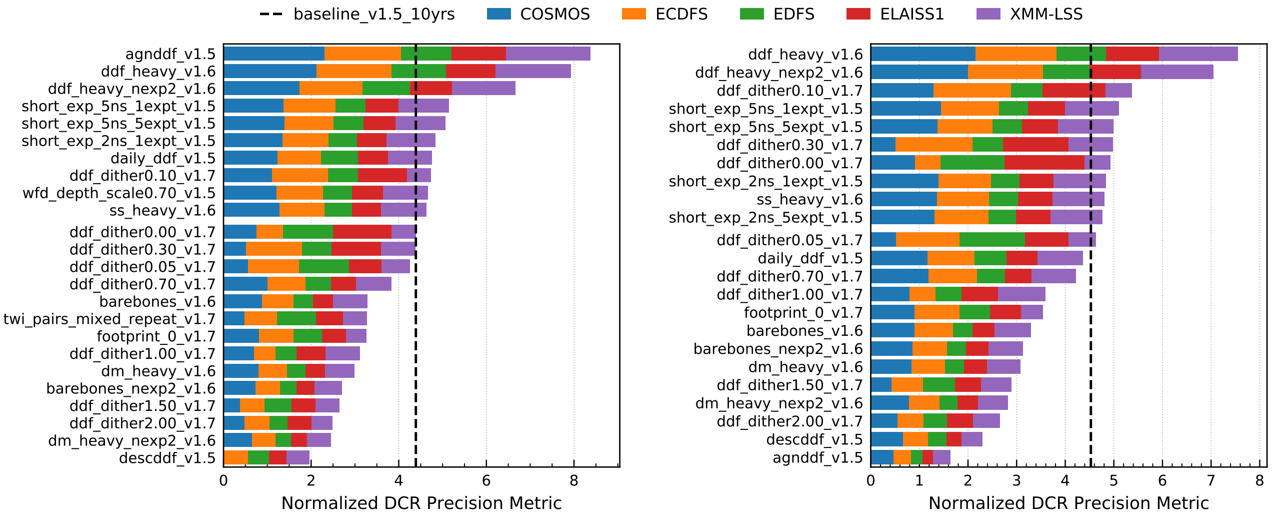

Figure 2. Relative improvement in DCR slope uncertainties within the five DDF fields (normalized to the "XMM-LSS" field in `baseline_v1.5_10yrs`; dashed vertical line) for the 10 best and 10 worst runs in addition to all OpSims that are designed specifically to test DDF survey strategies. The graph on the left shows the results obtained at $u = 22.15$ and the graph on the right shows the same evaluated at $g = 22$ —consistent with a typical $u - g = 0.15$ quasar color.

In the DDFs, the performance of each OpSim is jointly determined by its single-epoch depth (affecting astrometric precision), the number of visits, and the maximum airmass that the field is observed at. As expected, the `ddf_heavy` runs place at the top of the list. The `agnddf_v1.5` scores the highest in the u-band owing to its large number of visits and at the same time not being a strong negative outlier in the other two criteria. Three `short_exp` simulations make the top 10 in both u and g-band is a result of their slightly deeper single-epoch depth. OpSims that perform poorly overall also perform poorly in at least one of the three criteria. For example, `footprint_0_v1.7` is one of the worst in terms of the maximum airmass, as is `twi_paris_mixed_repeat_v1.7`. The `barebones` and `dm_heavy` simulations suffer from the relatively low number of visits. Lastly, *the* `descddf_v1.5` *and* `agnddf_v1.5` *rank* **last** *in the g-band primarily* *as a combined effect of their shallow single-epoch depth and low number of visits.* In addition, the issue related to lunar illumination in DDF scheduling as explored in (https://rtn-014.lsst.io) have played a role.

4. ANSWERS TO QUESTIONS

> **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.

Not applicable, but could be evaluated with the `footprint/footprint_tune` family simulations.

> **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.

We submitted a cadence whitepaper that suggested that a twilight survey might help to improve the utility of DCR. However, it seems that the number of epochs and the expected astrometric error are such that, even in the WFD, LSST is already providing suitable results for using DCR in the process of classification and photo-z. That said, a mini-survey that added high airmass observations in twilight (in u or g) should have the effect of improving the DCR signal. Existing twilight family OpSims do not (perhaps sensibly given the color of the sky) add observations in u or g and thus are not helpful for DCR.

> **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.

DCR in u band will benefit from longer exposure time if the number of visits do not drop dramatically, given that our DCR metric depends on the number of visits as well as the single epoch depth (which is correlated with the astrometric precision). As for the `u_long` simulations, the DCR slope uncertainty can be improved by nearly a factor of 1.5 with a 1x50 sec exposure (`u_long_ms_50_v1.7`) compared to the `baseline_nexp2_v1.7` run—without reducing the measurable DCR signal in the g-band (Yu et al. 2020b).

> **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.

DCR largely affects only the u and g bands. Except in extreme cases (e.g., fewer than 40 u-band observations), changes are unlikely to have a significant impact.

> **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.

Not simulated explicitly, but reducing the number of observations too low (see above), such that the number of high-airmass u or g-band observations is small, would have a negative impact on DCR. The `third_obs` simulations do not appear to be outliers (good or bad) in the DCR metric. Although the `baseline_samefilt_v1.5` performs slightly better than the baseline, we suspect that it is a direct result of an increased number of visits (Yu et al. 2020b).

> **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.

The `rolling` and `alt_rolling` simulations are not strong outliers (good or bad) in the DCR metric.

> **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.

From Figure 2, we can see that translational dithering with an offset larger than 1° can have a negative impact on DCR. The DCR slope uncertainty can increase by as much as a factor of 2 with a dithering offset of 2°. On the other hand, rotational dithers might help; see Ian Sullivan's work (https://dmtn-037.lsst.io/).

REFERENCES

Yu, W., Richards, G. T., Yoachim, P., & Peters, C. 2020a, Research Notes of the American Astronomical Society, 4, 252, doi: 10.3847/2515-5172/abd6e2

—. 2020b, doi: 10.17918/f5dn-8510