LSST AGN SC Cadence Note: Non-Parametric Structure Function Metric

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

1. EXECUTIVE SUMMARY

We have developed a model-independent metric ("SFErrorMetric") to assess the level to which we can derive the "structure function" of variable sources (e.g., AGNs) in LSST as a function of OpSims from versions FBS 1.5, 1.6 and 1.7. No presumptions about the actual underlying process that is responsible for the observed variability are used in this metric; this metric depends solely on the survey parameters (e.g., number of visits). Most of the survey simulations being considered for LSST operations performed equally well, with one exception being the `u_long` family, which significantly enhance this metric in the u-band without inducing observable drawbacks in other filters. Thus, we would favor longer u-band exposure time if the total number visits in the u-band can stay relatively unchanged.

2. THE METRIC

In the simplest form, the empirical structure function (SF) of a variable source at Δt can be derived from computing the arithmetic mean of the magnitude difference raised to the 2nd power of every pair of observations separated by Δt. Without knowing the true underlying process, we cannot determine the expected structure function (as to gauge any deviation and make comparisons), however, we can compute the uncertainty (variance) on the SF—modulo the portion depending on the *de facto* variability model. At each Δt bin, the variance on the SF can be expressed as follows:

$$\mathrm{Var}(\mathrm{SF}(\Delta \mathrm{t})) = 2 * \mu_{\sigma_i^2} + 2 * \sqrt{\mathrm{Var}(\sigma_i^2)/n} - 2 * cov(\Delta t) \tag{1}$$

which is obtained by propagating the photometric error on each observation through the structure function assuming small photometric errors. In Equation 1, σ_i is the photometric error, n is the number of pairs falling into the time bin centered at Δt, and $cov(\Delta t)$ is the auto-covariance of the underlying process at Δt. Since only the last term on the right-hand side of Equation 1 depends on the true underlying variability structure, which is independent of the survey parameters, differential comparisons of OpSims can be made by simply comparing the sum of the first two terms on the right-hand side of Equation 1 at each Δt bin, provided that the number of bins used in evaluating this metric is consistent across all OpSims. An overall metric can be acquired by adding up the SF uncertainties over all Δt bins. The current implementation of this metric uses 15 Δt bins that are uniformly spaced in the log scale (ranging from 0.01 day to 10 years) for the Wide-Fast-Deep (WFD) survey and 20 such Δt bins for the Deep-Drilling Fields (DDFs) (Yu & Richards 2021).

In Figure 1, we present the results obtained in the WFD as distributions of relative improvements in the metric with respect to the `baseline_v1.5_10yrs` simulation. The top row of Figure 1 displays the metric results evaluated at $u = 24.15$ and the bottom row shows the results from $r = 23.85$ (matching the projected 5σ single-epoch depth in WFD and consistent with a typical $u - r = 0.3$ quasar color). The two panels on the left plot the distributions for one member of each OpSim "family" (with the best metric for that family) and 5 baseline OpSims (that are used as comparison runs). Families that specifically test strategies in the DDFs are not included. The two panels on the right show the results for 10 best and 10 worst OpSims out of all relevant runs.

Shown in Figure 2 is the median metric at each DDF with the metric 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 OpSims plotted here are the 10 best and 10 worst OpSims as well as all DDF-specific experiments (excluding those in the `euclid_dithers` family unless they appear to be the best or worst). The left and right panels show the results obtained in the u and r filters, respectively. Note that the normalized metric only indicates relative improvement in the part of the variance on the SF due to survey parameters, rather than the improvement in the full uncertainty (which is inevitably model-dependent).

3. ANALYSIS OF RESULTS

In the WFD survey, the largest variance in this structure function metric was seen in the u-band, where the `u_long` simulations performed much better than the rest (see the top row of Figure 1). Other simulations that also scored high are basically comparable to the baseline run. Eight out of 10 OpSims that performed the poorest are variants

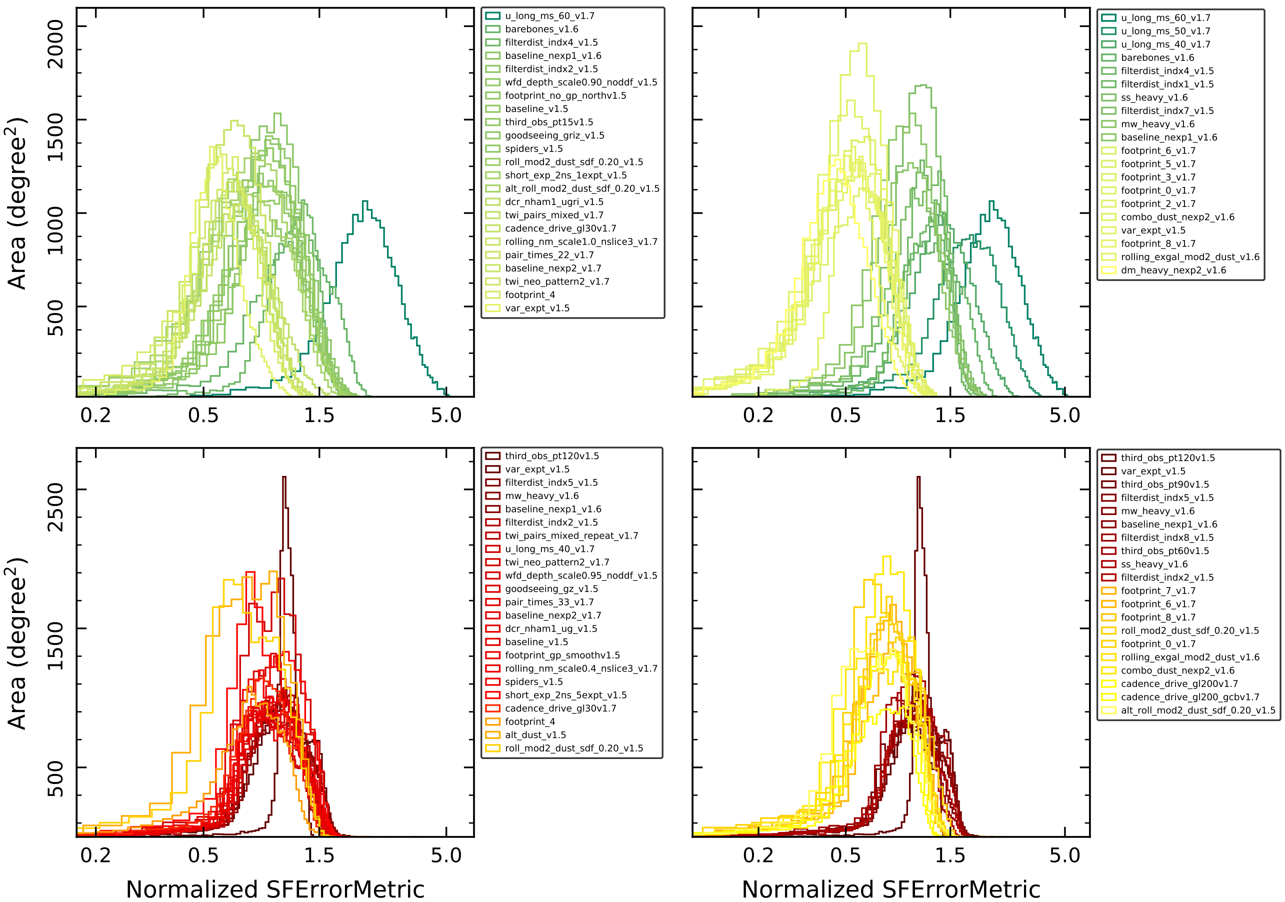

Figure 1. Distribution of relative improvement in the structure function metric (on a log axis and normalized to the median metric of the `baseline_v1.5_10yrs` simulation); larger is better. *Top Left:* Shown are a total of 23 different LSST OpSims including one (with the best metric) from each family that vary the survey strategy in the WFD at a fiducial depth of $u = 24.15$. *Top Right:* The same distributions at $u = 24.15$ are plotted for the 10 best and 10 worst OpSims of all relevant OpSims. *Bottom Left/Right:* Same as the two panels in the top row, but showing the results evaluated at $r = 23.85$.

of the `combo_dust` simulation (v1.6) including 6 versions from the `footprint_tune` family. We suspect that this poor performance is a combined effect of a reduced number of visits and a shallower single-epoch depth. Also scoring poorly is the `var_exp_v1.5` simulation, which has a much shallower single-epoch depth (≈ 0.6 mag shallower). That the `dm_heavy_nexp2_v1.6` simulation ranks last is likely due to the increased variance in the distribution of photometric errors (at a fixed magnitude) as a result of an increased number of high airmass visits.

In the r-band, most of the OpSims that performed the best provide only marginal improvement over the baseline, where the `var_expt_v1.5` is worth more attention given that it appeared as one of the worst in the u-band. We believe the reason behind this disparity is that forcing a uniform depth has a relatively milder impact on the single-epoch depth in the r-band (≈ 0.2 mag drop) compared to that in the u-band (≈ 0.6 mag drop). Among the ones that performed the worst, 5 out of 10 are variants of the `combo_dust` simulation and three more with rolling strategy inserted. The remaining two are part of the `cadence_drive` experiment, which scored poorly due to a combination of a slightly shallower single-epoch depth and a reduced number of visits. Although not shown explicitly in this note, the `cadence_drive` experiments impose a strong negative effect in the g-band as the fill-in visits (took under less photometric conditions) drive up the mean and variance of the distribution of the photometric errors (see Yu & Richards 2021). In addition to what is being shown in Figure 1, the `dcr` simulations introduced up to a $\approx 40\%$ drop of this metric in u and g filters. Likewise, the median metrics for the `short_exp` runs are up to $\approx 20\%$ lower in those two filters.

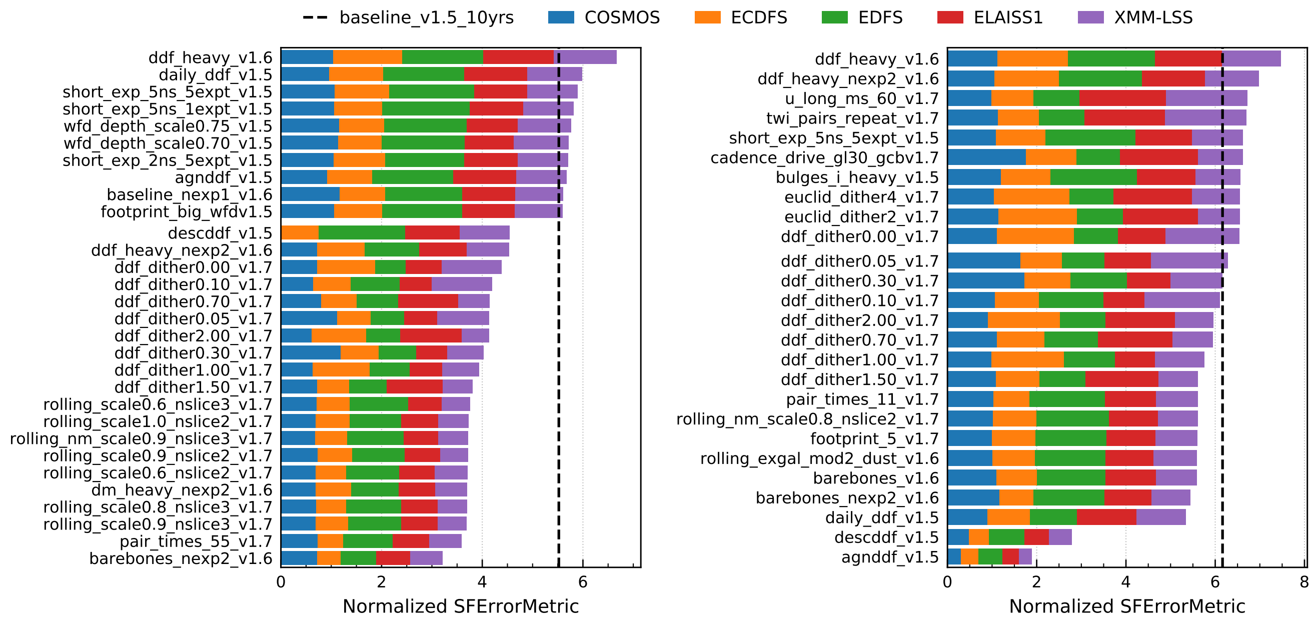

Figure 2. Relative improvement in the structure function metric within the five DDF fields (normalized to the "XMM-LSS" field in `baseline_v1.5_10yrs`; dashed vertical line) for the 10 best and 10 worst runs in addition to the OpSims that are designed specifically to test DDF survey strategies (excluding those in the `euclid_dithers` family if not among the best or worst). The graph on the left shows the results obtained at $u = 24.15$, and the graph on the right shows the same evaluated at $r = 23.85$—consistent with a typical $u - r = 0.3$ quasar color.

From Figure 2 we can see that there is not as much variance in this metric in the DDFs as in the WFD. The variance seen in the u-band (left panel) is mostly driven by the variance in the single-epoch depth and the number of visits. In the r-band (right panel), we identify two strong outliers: `descddf_v1.5` and `agnddf_v1.5`. Further investigation revealed that a large portion of the r-band visits were carried out at a much higher sky brightness (than the nominal WFD visits), resulting in a 1 magnitude drop (on average) in the single-epoch depth—in addition to the expected less frequent re-visits in the g and r bands in those two OpSims. This observed effect (high sky brightness) has been explored in much more detail in a Rubin Observatory technote, RTN-014: "Lunar Complications in the Scheduling of Deep Drilling Fields" (Neilsen et al. 2021). *Future simulations should consider ways to minimize such effect as much as possible while trying to avoid long gaps in the DDF sequences.*

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.

Per the SFErrorMetric, variations in the footprint do not strongly influence the uncertainty level on the structure function. That said, a $\approx 15\%$ drop in performance relative to the corresponding baseline is not unusual. We suspect that this is a direct result of the reduced number of visits and shallower single-epoch depth (as seen in the `footprint_tune` runs). However, we stress that further reducing the number of visits in the u-band could result in even more empty Δt bins (zero pair of visits found around Δt) in the process of computing the structure function. Given a median of ≈ 56 u-band visits over 10 years (as in the current baseline) and 15 uniformly spaced Δt bins in the log scale, a 10% drop in the number of visits will introduce $\approx 1\%$ more empty bins. Thus, we would prefer the number of visits in the u-band to remain unaffected.

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.

Not applicable.

> **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.

Longer u-band exposures without a significantly reduced number of visits (e.g., halving) will benefit structure function analysis in the u-band. The SFErrorMetric is enhanced by nearly a factor 3 in the u-band (at a fixed magnitude) while changing the exposure time from 2x15 sec (`baseline_nexp2_v1.7`) to 1x50 sec (`u_long_ms_50_v1.7`), and no recognizable drawbacks were identified in either g or r-band.

> **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.

Influenced by cosmological time dilation effects and the fact that AGN variability amplitude reduces toward longer wavelength, AGN structure function analysis could benefit from more visits in bluer filters.

> **Q5**: Are there any science drivers that would strongly argue for, or against, obtained 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.

Given the simulations that are currently available, pair strategies in the WFD survey do not seem to have any strong impact on structure function analysis. However, obtaining two visits in the same filter (`baseline_samefilt_v1.5`) can introduce some drawbacks.

> **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.

Extreme rolling cadence (all visits in one year) will be very detrimental to AGN variability science. Among the current realizations of the rolling strategy, the versions without nightly modulation are preferred. The rolling simulations in the other class (with nightly modulation) are slightly worse ($\approx 8\%$ drop in the normalized metric) than the corresponding baseline in both u and g bands; a $\approx 15\%$ drop was observed in the r-band.

> **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.

The exact dithering pattern does not seem to have any noticeable strong effect on the structure function uncertainty in, e.g., the DDFs.

REFERENCES

Neilsen, E., Jones, L., & Yoachim, P. 2021. https://rtn-014.lsst.io

Yu, W., & Richards, G. T. 2021. https://github.com/RichardsGroup/LSST_SF_Metric/