

Testing of LSST AGN Selection Using Rolling Cadences

Gordon Richards (Drexel), Weixiang Yu (Drexel), W.N. Brandt (Penn State),
Qingling Ni (Penn State), Christina Peters (Toronto), Guang Yang (Penn State),
Franz E. Bauer (PUC/SSI)

November 2018

Abstract

The Supernova community has expressed an interest in so-called “rolling cadences”, in order to achieve better sampling of SNe light curves. Here we explore how such rolling cadences would affect both the identification of AGNs and the science that can be derived from their resulting light curves. We present a metric that quantifies how well parameters of a particular variability model would be recovered for a given opSim cadence. We conclude that extreme rolling cadences (with few or no observations during some seasons) would be detrimental to AGN science. Less extreme rolling cadences require further investigation as the answer depends critically on both the science and the desired cadence.

1 White Paper Information

- Contact author: Gordon Richards (Drexel University), gtr@physics.drexel.edu
- **Science Category:** This white paper addresses issues related to recognizing and characterizing AGNs via their variability properties and is thus most relevant to “Exploring the Changing Sky”.
- **Survey Type Category:** Primarily related to extending the benefit of the Wide-Fast-Deep (WFD) survey area.
- **Observing Strategy Category:** This proposal is related to the observing strategy needed to enable specific WFD time-domain science. The area and location of the WFD are not addressed.

2 Scientific Motivation

This white paper explores the question of what rolling cadence¹ choices would work acceptably well for AGN science in the WFD region. The context is both in terms of *identifying* AGNs (as variable sources with characteristic variability parameters) and *characterizing* AGNs (specifically the reliability with which the AGN light curves are represented by a model) as related to the ability to do AGN-related science within a given cadence scheme.

To put the problem/question in context, it helps to think about the differences between the light curves of an AGN and a Supernova. For SNe, the light curves are ordered in time: the goal is to see a pattern that matches a fixed template where all of the action occurs over a period of ~ 180 days. Thus it is crucial for SNe to set the cadence to something that fully samples the light curve *during that 180 day period*.

AGNs, on the other hand, have stochastic light curves, that can be well-represented by a so-called “damped random walk (DRW)” (MacLeod et al., 2010; Moreno et al., 2018). In the case of AGNs, the light curve can be characterized as “stationary”, meaning that any segment of the light curve is qualitatively similar to any other segment of the light curve (of the same length). Thus there is no specific need to concentrate observations in time.

That said, the sampling cadence does very much matter for AGNs as we will demonstrate. Some rolling cadences would clearly be detrimental for AGN science; however, it may well be that certain rolling cadences would end up being beneficial for AGN science too. The impact of identifying and characterizing AGN may extend to other extragalactic science, through the improvement of photometric redshifts and/or the identification of outliers. Thus the point of this white paper is to emphasize that the ramifications of adopting a significantly non-uniform cadence should be tested in light of AGN science. We provide one simple metric using the recovery of an input AGN model (which we suggest is also correlated with the completeness/efficiency of AGN selection) as a way to assess the impact of the cadence on AGNs.

The empirical goals that ultimately underlie AGN science are three-fold with respect to the question of cadence in the WFD:

1. Determine if each object on the sky is consistent with being a variable source.
2. Determine if said variability is consistent with the source being an AGN.
3. Determine model parameters that can be used to probe AGN physics.

The goal is therefore *not* to fit a specific model template (as it is for SNe) and, even in the case of item 3, the goal is to fit a *family* of models and determine the best-fit coefficients. Those coefficients then inform the AGN selection criteria and are the foundation for using AGNs to understand accretion disk physics.

¹Defined here as cadences that concentrate observing epochs into a few (sometimes repeating) observing seasons instead of spreading them uniformly over the lifetime of the survey.

These goals lead to potentially very different cadence requirements for SNe and AGNs. For AGNs, rather than concentrating cadences in time to fit a model, AGN science would prefer that the cadences be as close to uniform in $\log \Delta t$ —rather than uniform in time. We illustrate this in Figure 1 where we compare the “structure function” (sorting of the light curve into all possible Δt bins) of an AGN—sampled first for a uniform cadence and then for a rolling cadence. Both exhibit aliasing (non-uniformity in sampling with some Δt values having many more samples than others), but the rolling cadence has stronger aliasing than the uniform cadence. Specifically, certain time separations have been sampled at a rate that exceeds the average sampling rate for all time separations. For example, while the rolling cadence *does* improve the sampling at small Δt (by definition), the aliasing at $365 * 3 = 1095$ days (for a 3-year rolling period) is inefficient. There are already enough observations at $\Delta t = 1095$ days; it would be preferable from the standpoint of fitting AGN models to light curves to instead conduct observations that would help to fill in a gap in Δt space.

It may be that AGNs could actually *benefit* from the adoption of some sort of rolling cadence. For example, concentrating observations in one season will increase the number of observations at small Δt . Concentrations in further seasons could help probe short timescale variability and, if those additional periods of high cadence are not distributed in a fixed pattern (e.g., a three-year cycle), they could further help fill holes in the Δt distribution. In addition, the desire to do “photometric reverberation mapping (PRM)” of the accretion disk and broad line region, which involves cross-correlating the light curves from at least two photometric bands, will benefit from an appropriate rolling cadence (e.g., Chelouche et al. 2014; Fausnaugh et al. 2016; Jiang et al. 2017; Edelson et al. 2018; Mudd et al. 2018; Yu et al. 2018). Simply put, a higher sampling rate might make it easier to identify “features” in the light curves that would help to constrain the differences in characteristic timescales (and thus distances) between light curves from different bands.

Below we compare results for simulated quasars using 4 different opSim observing cadences. As each opSim is unique and the science goals of AGN identification and characterization (including PRM) may be optimized for different cadences, it is difficult to give concrete general conclusions about the impact of rolling cadences. However, the following conclusions should be robust: 1) For any cadence choice, longer observing seasons help to reduce aliasing of the time sampling; 2) a *single season* of high-cadence observations (in any given area of sky) would reduce aliasing compared to a repeating pattern (while increasing the sampling at small Δt); 3) fixed-pattern repeated-year rolling cadences will lead to (undesirable) aliasing in Δt .

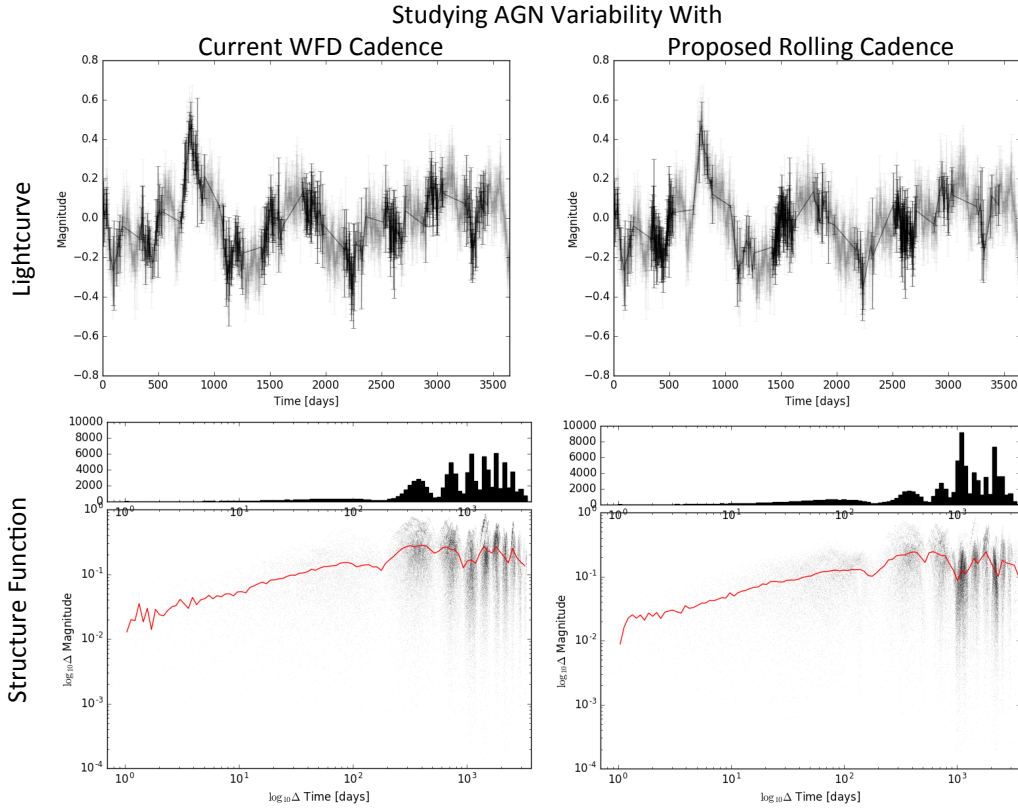


Figure 1: (Top panels): Simulated AGN light curve sampled daily (grey). On the left we illustrate 10 observing seasons of uniform cadence (black). On the right we illustrate a rolling cadence with preferential clustering of observations in years 2, 5, and 8. The middle panels present a histogram of all the time separations between observations. Significantly stronger (undesirable) aliasing is observed in the right-hand histogram at ~ 1095 days (three years) where many observations in years 2, 5, and 8 are contributing. AGN structure functions would be best probed by sampling that produces a more uniform histogram in the middle panels (i.e., uniform in $\log \Delta t$). The best-fit structure functions are shown in the bottom panels.

3 Technical Description

Describe your survey strategy modifications or proposed observations. Please comment on each observing constraint below, including the technical motivation behind any constraints. Where relevant, indicate if the constraint applies to all requested observations or a specific subset. Please note which constraints are not relevant or important for your science goals.

3.1 High-level description

Describe or illustrate your ideal sequence of observations.

Simply put, we desire an observing cadence that is as uniform as possible (given seasonal gaps) in $\log \Delta t$. Clustering of observations in a single season would yield a desirable sampling at small Δt (that a uniform cadence would not provide), but repeating clustering in time of the same area of sky leads to undesirable aliasing.

Within any choice of cadence, one way to make the the Δt sampling more uniform is to have observing seasons that are as long as possible.

We also require that reasonable-depth observations be maintained over the full 10 years of the LSST survey. This will let variability studies reach the longest possible timescales, where the red-noise power spectrum of AGN variability implies that variability will be the strongest—this will aid AGN selection using variability. Reaching the longest possible timescales will also provide the best constraints on proper motion, again aiding AGN selection.

3.2 Footprint – pointings, regions and/or constraints

Describe the specific pointings or general region (RA/Dec, Galactic longitude/latitude or Ecliptic longitude/latitude) for the observations. Please describe any additional requirements, especially if there are no specific constraints on the pointings (e.g. stellar density, galactic dust extinction).

Nominal WFD

3.3 Image quality

Constraints on the image quality (seeing).

Nominal WFD

3.4 Individual image depth and/or sky brightness

Constraints on the sky brightness in each image and/or individual image depth for point sources. Please differentiate between motivation for a desired sky brightness or individual image depth (as calculated for point sources). Please provide sky brightness or image depth constraints per filter.

Nominal WFD

3.5 Co-added image depth and/or total number of visits

Constraints on the total co-added depth and/or total number of visits. Please differentiate between motivations for a given co-added depth and total number of visits. Please provide desired co-added depth and/or total number of visits per filter, if relevant.

Nominal WFD

3.6 Number of visits within a night

Constraints on the number of exposures (or visits) in a night, especially if considering sequences of visits.

Nominal WFD

3.7 Distribution of visits over time

Constraints on the timing of visits — within a night, between nights, between seasons or between years (which could be relevant for rolling cadence choices in the WideFastDeep). Please describe optimum visit timing as well as acceptable limits on visit timing, and options in case of missed visits (due to weather, etc.). If this timing should include particular sequences of filters, please describe.

We desire an observing cadence that is as uniform as possible (given seasonal gaps) in $\log \Delta t$. Potential sources of concern with regard to the nominal cadence or other proposals are: repeat clustering of observations in the same area of sky, short season lengths, etc.

Again, observations should be distributed over the full 10-year lifetime of the LSST survey.

3.8 Filter choice

Please describe any filter constraints not included above.

Although this white paper does not specifically address the integration of multiple band-passes into AGN variability investigations, we emphasize that there are *crucial* filter constraints for AGN science. The *u*-band is by far the most important at $z < 3$ and the *z* and *y*-bands for $z > 6$. Filter-balance considerations fundamentally interact with the number of epochs that can be obtained, given the global constraint on total possible LSST visits over 10 years.

From the standpoint of variability analysis it is important to maximize the number of epochs in *each* filter as the wavelength dependence of variability provides constraints on accretion disk physics. It is further desirable (e.g., for the accuracy of photometric redshifts) to achieve similar depths in each of the filters; an example of filter balance that maintains similar depths in each of the filters is given in Section 3.4 of Brandt et al. (arXiv:1811.06542).

3.9 Exposure constraints

Describe any constraints on the minimum or maximum exposure time per visit required (or alternatively, saturation limits). Please comment on any constraints on the number of exposures in a visit.

Nominal WFD

3.10 Other constraints

Any other constraints.

3.11 Estimated time requirement

Approximate total time requested for these observations, using the guidelines available at https://github.com/lstt-pst/survey_strategy_wp.

Nominal WFD

Properties	Importance
Image quality	3
Sky brightness	3
Individual image depth	3
Co-added image depth	3
Number of exposures in a visit	2
Number of visits (in a night)	2
Total number of visits	1
Time between visits (in a night)	2
Time between visits (between nights)	1
Long-term gaps between visits	1
Other (please add other constraints as needed)	

Table 1: **Constraint Rankings:** Summary of the relative importance of various survey strategy constraints. Please rank the importance of each of these considerations, from 1=very important, 2=somewhat important, 3=not important. If a given constraint depends on other parameters in the table, but these other parameters are not important in themselves, please only mark the final constraint as important. For example, individual image depth depends on image quality, sky brightness, and number of exposures in a visit; if your science depends on the individual image depth but not directly on the other parameters, individual image depth would be ‘1’ and the other parameters could be marked as ‘3’, giving us the most flexibility when determining the composition of a visit, for example.

3.12 Technical trades

To aid in attempts to combine this proposed survey modification with others, please address the following questions:

1. *What is the effect of a trade-off between your requested survey footprint (area) and requested co-added depth or number of visits?*
2. *If not requesting a specific timing of visits, what is the effect of a trade-off between the uniformity of observations and the frequency of observations in time? e.g. a ‘rolling cadence’ increases the frequency of visits during a short time period at the cost of fewer visits the rest of the time, making the overall sampling less uniform.*
3. *What is the effect of a trade-off on the exposure time and number of visits (e.g. increasing the individual image depth but decreasing the overall number of visits)?*
4. *What is the effect of a trade-off between uniformity in number of visits and co-added depth? Is there any benefit to real-time exposure time optimization to obtain nearly constant single-visit limiting depth?*
5. *Are there any other potential trade-offs to consider when attempting to balance this proposal with others which may have similar but slightly different requests?*

It is not the goal of this white paper to propose a specific cadence for LSST as it is not clear that AGNs should be the driver of that decision. Our purpose is simply to point out that adoption of a rolling cadence may have unintended consequences (“trade-offs”) and that the consequences for AGNs should be weighed in the decision.

4 Performance Evaluation

We are currently evaluating the utility of different opSims by determining how well known parameters of DRW models of AGNs are recovered (given a realistic suite of DRW parameter space). We illustrate this below and present a possible metric.² Currently the computation of the metric requires our *Kali* package (Kasliwal; <https://github.com/AstroVPK/kali>) that does MCMC fitting of AGN light curves. This metric primarily focuses on the 3rd goal indicated above (related to AGN science) and not the first 2 (related to AGN discovery), but the ability to accurately constrain model parameters is certainly correlated with the ability to use those model parameters for object classification.

We first draw a grid over the DRW parameter space (characteristic timescale, τ , and the long-term amplitude of variability, SF_∞) occupied by SDSS Stripe 82 quasars. This grid is indicated by colored stars in Figure 2. For each point on the grid, we simulate an AGN light curve (with 30-minute cadence) using corresponding model parameters. Then we downsample each full light curve according to observations from a given opSim. Test “objects” are drawn from positions evenly spaced (5 degrees along Declination and 10 degrees along Right Ascension) across the Wide-Fast-Deep footprint of LSST. By fitting the downsampled light curves to the DRW model and plotting the best-fit parameters in the DRW parameter space, we generate the clouds as shown in Figure 2.

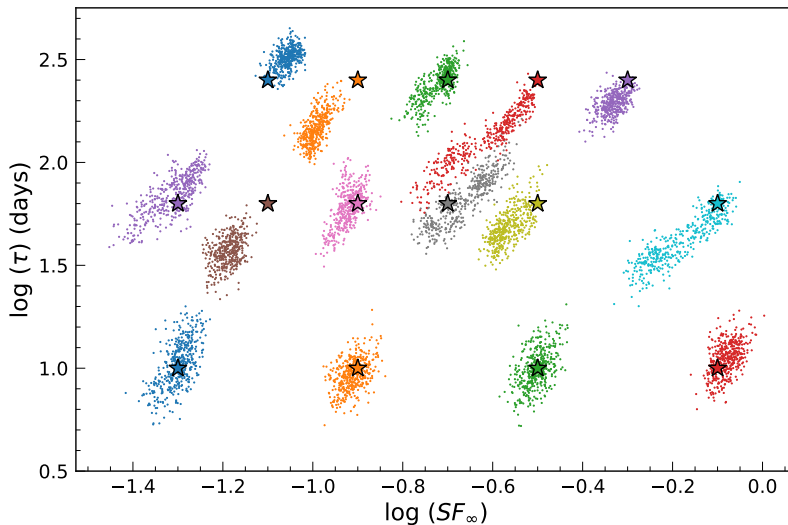


Figure 2: The stars represent the input parameters for the simulated AGNs and the clouds of points are corresponding best-fit parameters of each downsampled light curve. The best-fits and the inputs are matched by colors.

Two metrics capture the reliability with which the known model parameters are recovered: a ‘shift’ of the cloud centroid and the ‘width’ of the cloud. We define the shift of the cloud

²The core code is available at <https://github.com/RichardsGroup/LSSTprep/blob/master/Rolling/Metric.ipynb>

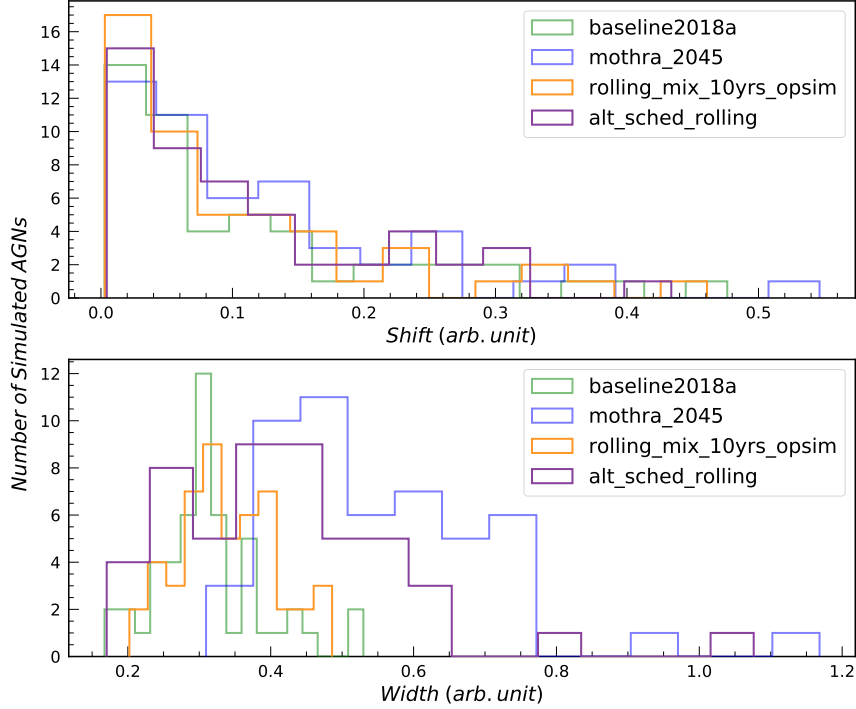


Figure 3: Top: The distribution of the shift from the input parameters to the center of the best-fit parameter cloud for fifty simulated AGNs. Bottom: The distribution of the width of the best-fit parameter cloud for the same set of simulated AGNs. Large width indicates an opSim that may be detrimental to characterizing/selecting AGNs in the WFD.

from its input by the Euclidean distance from the center of the cloud to the input, where the center is determined by the median of the best-fit parameters in that cloud. As biases in shift can be modeled out, the ‘width’ of each cloud is the preferred metric and is defined here as the $2\text{-}\sigma$ width along the semi-major axis. The distributions of shift and width for the best-fit parameter clouds generated by light curves downsampled from four different simulated cadences are shown in Figure 3.

We find that the `mothra_2045` opSim (which adopts a rolling cadence) shows considerably more scatter than the `baseline2018a` opSim, suggesting that `mothra_2045` would be bad for AGN science (and likely bad for AGN selection). However, that opSim apparently has a bug. Instead `rolling_mix_10yrs_opsim` is supposed to be a more appropriate rolling cadence opSim, and it appear to be less problematic. The result from the `alt_sched_rolling`, simulated by the LSST DESC, also shows similar scatter as seen in the `mothra_2045` opSim.

What `mothra_2045` and `alt_sched_rolling` have in common is that they completely turn off visits in the deprioritized regions, whereas the `rolling_mix_10yrs_opsim` maintains 25% of the baseline visits in the deprioritized regions. Therefore, we stress that the less extreme version of the rolling cadences is *strongly* preferred for AGN science.

5 Special Data Processing

Describe any data processing requirements beyond the standard LSST Data Management pipelines and how these will be achieved.

No special processing required.

References

Chelouche, D., Shemmer, O., Cotlier, G. I., Barth, A. J., & Rafter, S. E. 2014, , 785, 140

MacLeod, C. L., Ivezić, Ž., Kochanek, C. S., et al. 2010, , 721, 1014

Moreno, J., Vogeley, M. S., & Richards, G. T. 2018, ArXiv e-prints, [arXiv:1811.00154](https://arxiv.org/abs/1811.00154) [astro-ph.IM]