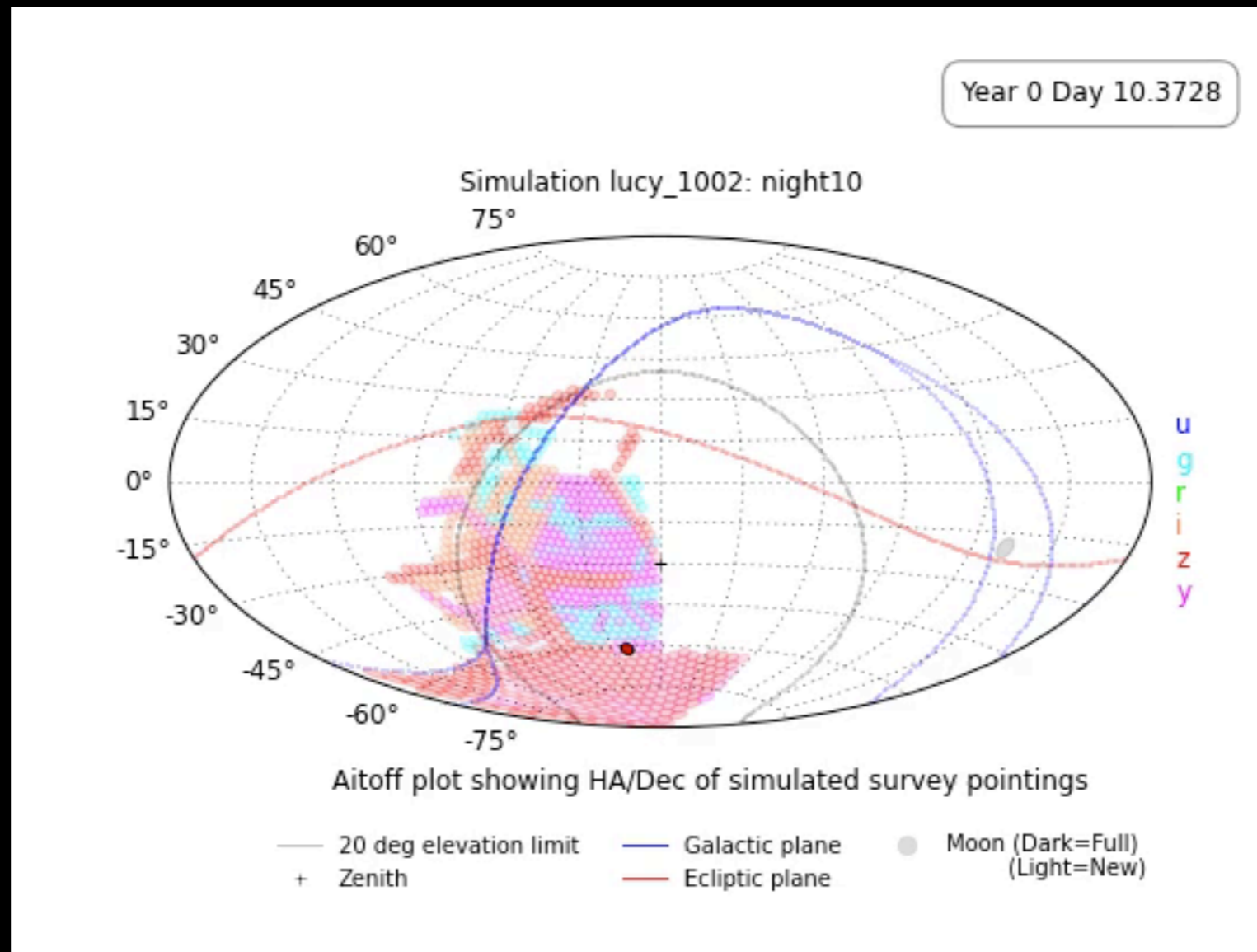


METRICS ANALYSIS FRAMEWORK

# OVERVIEW OF SCHEDULER METRICS

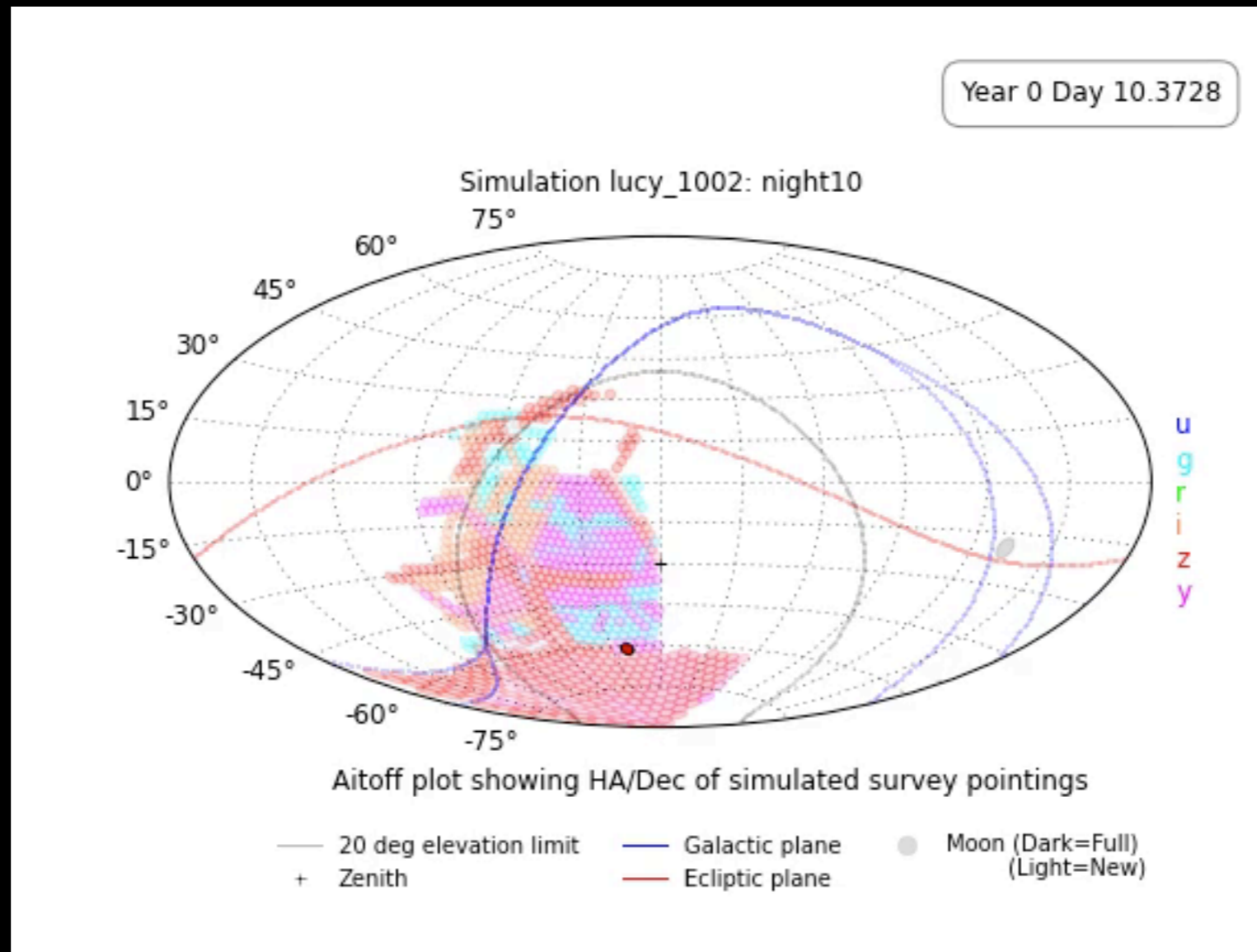
LYNNE JONES @ MARCH 2015 SCHEDULER WORKSHOP

- Scheduler generates a simulated pointing history for LSST



- How do we know if the observing strategies and scheduling choices that generated that pointing history were optimal (or even 'good enough')?

- Scheduler generates a simulated pointing history for LSST



- How do we know if the observing strategies and scheduling choices that generated that pointing history were optimal (or even 'good enough')?

# EVALUATING THE SCHEDULER: METRICS

- **Science metrics**

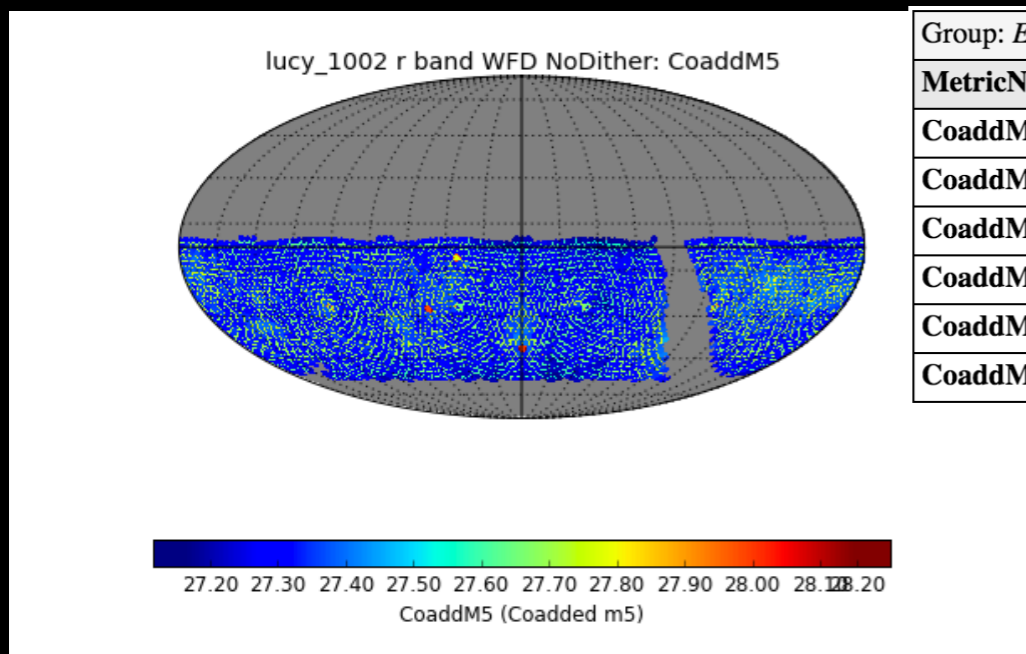
- Evaluate scientific performance of scheduler
- Does survey meet minimum/design/stretch science requirements laid out in Science Requirements Document (SRD)?
- How well does survey perform in a wide variety of other science goals?

- **Technical metrics**

- Evaluate technical performance of scheduler
- Were requested visits obtained? In desired conditions? How efficiently did scheduler obtain visits?

# METRICS ANALYSIS FRAMEWORK (MAF)

- MAF is an open-source python framework designed to evaluate simulated surveys (scheduler outputs)
  - Flexible - apply customizable set of metrics
  - Extensible - easily generate new metrics
  - Evaluate metrics across the sky on a variety of spatial scales (e.g. healpixel grids) or on other subdivisions of pointing history (e.g. airmass or seeing grid) and generate visualizations + summary statistics

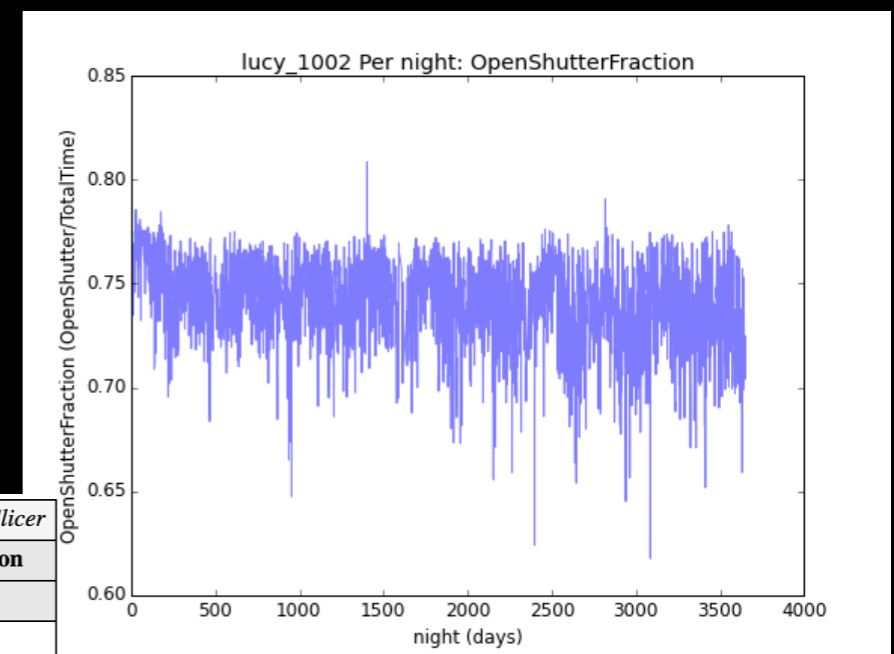


Group: *E: Coadded depth*; Subgroup: *All Props*; Slicer: *Op*

MetricName	Metadata	Median	Mean	Rms
CoaddM5	u band, all props	26.05	25.93	0.37
CoaddM5	g band, all props	27.13	26.98	0.34
CoaddM5	r band, all props	27.29	27.03	0.54
CoaddM5	i band, all props	26.54	26.35	0.48
CoaddM5	z band, all props	25.25	25.17	0.28
CoaddM5	y band, all props	24.30	24.08	0.49

Group: *A: Summary*; Subgroup: *3: Obs Per Night*; Slicer: *OneDSlicer*

	Filter Changes	NVisits	OpenShutterFraction
	Per night	Per night	Per night
Median	2.00	824.50	0.74
Mean	3.92	814.63	0.74
Rms	3.92	185.28	0.0205



# EVALUATING THE SCHEDULER: METRICS

- **Science metrics**

- Evaluate scientific performance of scheduler
- Does survey meet minimum/design/stretch science requirements laid out in Science Requirements Document (SRD)?
- How well does survey perform in a wide variety of other science goals?

- **Technical metrics**

- Evaluate technical performance of scheduler
- Were requested visits obtained? In desired conditions? How efficiently did scheduler obtain visits?

# SCIENCE METRICS: SRD

- Science Requirements Document (SRD) - LPM-17
- Intentionally provides limited constraints on cadence
- Specifications flow from four primary science drivers for LSST and provide minimum/design/stretch goals
  - Failing minimum requirements implies significant problems addressing main science drivers (as well as other science)

# SCIENCE METRICS: SRD

- Sky area uniformly covered by the main (WFD) survey
- Median number of visits in all bands across this sky area
- Sky area with nearly uniform visits on short timescales, 40s-30minutes
- Median trigonometric parallax accuracy for  $r=24$  unresolved sources

**Specification:** The sky area uniformly covered by the main survey will include Asky square degrees (Table 22).

Quantity	Design Spec	Minimum Spec	Stretch Goal
Asky (deg <sup>2</sup> )	18,000	15,000	20,000

Table 22: The sky area uniformly covered by the main survey.

**Specification:** The sum of the median number of visits in each band,  $N_{v1}$ , across the sky area specified in Table 22, will not be smaller than  $N_{v1}$  (Table 23).

Quantity	Design Spec	Minimum Spec	Stretch Goal
$N_{v1}$	825	750	1000

Table 23: The sum of the median number of visits in each band across the sky area specified in Table 22.

**Specification:** At least  $RVA1$  square degrees will have multiple observations separated by nearly uniformly sampled time scales ranging from 40 sec to 30 min (Table 25).

Quantity	Design Spec	Minimum Spec	Stretch Goal
$RVA1$ (deg <sup>2</sup> )	2,000	1,000	3,000

Table 25: The minimum area with fast (40 sec – 30 min) revisits.

**Specification:** The median trigonometric parallax accuracy across the main survey area for unresolved sources with  $r = 24$  must be at least  $SIG_{para}$  mas. The median proper motion accuracy per coordinate across the main survey area for such sources must be at least  $SIG_{pm}$ . The median trigonometric parallax accuracy across the main survey area for unresolved sources detected only in the  $y$  band (at  $10\sigma$ ) must be at least  $SIG_{paraRed}$  mas. (Table 26).

Quantity	Design Spec	Minimum Spec	Stretch Goal
$SIG_{para}$ (mas)	3.0	6.0	1.5
$SIG_{pm}$ (mas/yr)	1.0	2.0	0.5
$SIG_{paraRed}$ (mas)	6.0	10.0	3.0

Table 26: The required trigonometric parallax and proper motion accuracy.



# SCIENCE METRICS: SRD

- Sky area uniformly covered by the main (WFD) survey
- Median number of visits in all bands across this sky area
- Sky area with nearly uniform visits on short timescales, 40s-30minutes
- Median trigonometric parallax accuracy for  $r=24$  unresolved sources

**Specification:** The sky area uniformly covered by the main survey will include Asky square degrees (Table 22).

Quantity	Design Spec	Minimum Spec	Stretch Goal
Asky (deg <sup>2</sup> )	18,000	15,000	20,000

Table 22: The sky area uniformly covered by the main survey.

**Specification:** The sum of the median number of visits in each band,  $N_{v1}$ , across the sky area specified in Table 22, will not be smaller than  $N_{v1}$  (Table 23).

Quantity	Design Spec	Minimum Spec	Stretch Goal
$N_{v1}$	825	750	1000

Table 23: The sum of the median number of visits in each band across the sky area specified in Table 22.

**Specification:** At least RVA1 square degrees will have multiple observations separated by nearly uniformly sampled time scales ranging from 40 sec to 30 min (Table 25).

Quantity	Design Spec	Minimum Spec	Stretch Goal
RVA1 (deg <sup>2</sup> )	2,000	1,000	3,000

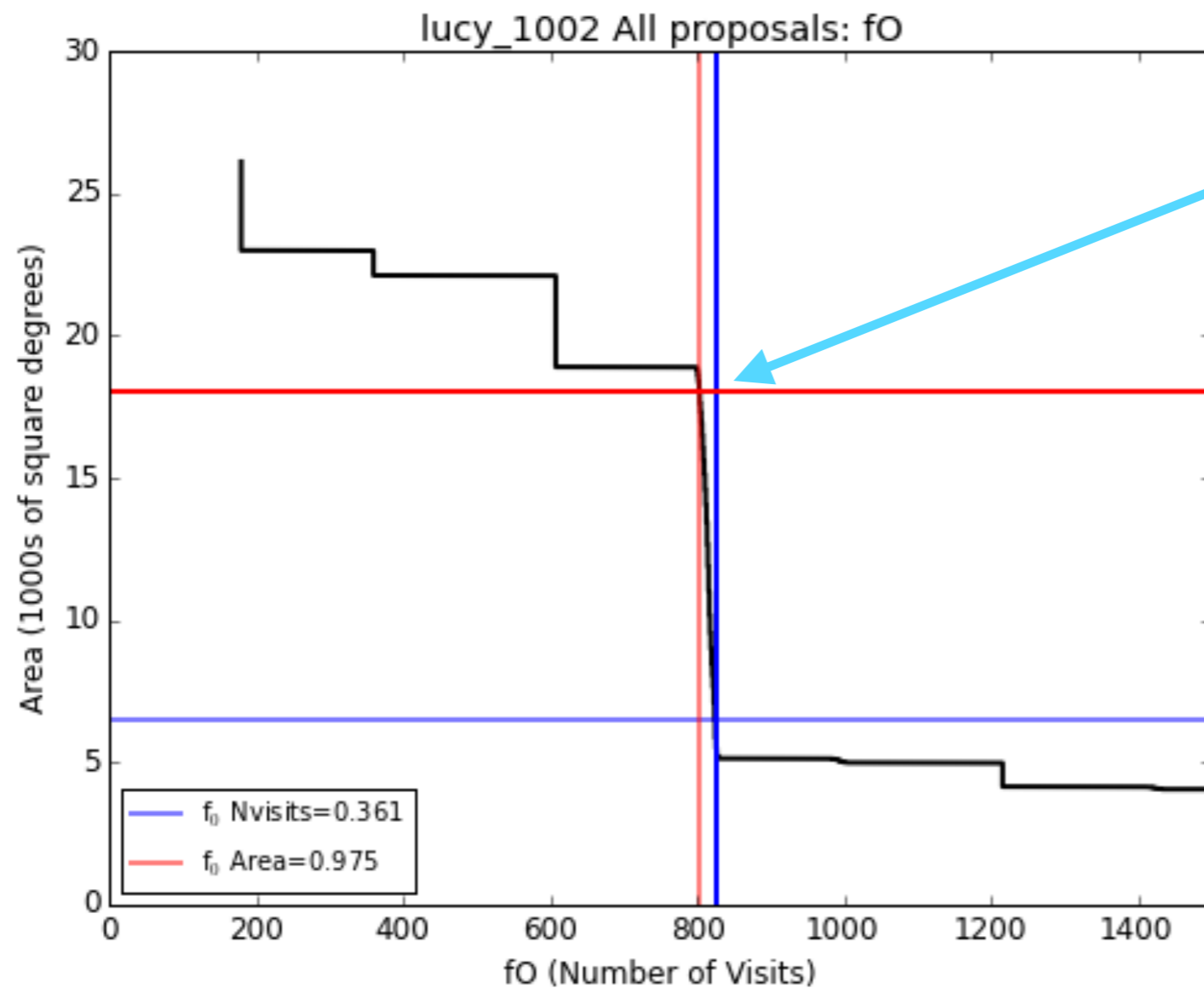
Table 25: The minimum area with fast (40 sec – 30 min) revisits.

**Specification:** The median trigonometric parallax accuracy across the main survey area for unresolved sources with  $r = 24$  must be at least SIGpara mas. The median proper motion accuracy per coordinate across the main survey area for such sources must be at least SIGpm. The median trigonometric parallax accuracy across the main survey area for unresolved sources detected only in the  $y$  band (at  $10\sigma$ ) must be at least SIGparaRed mas. (Table 26).

Quantity	Design Spec	Minimum Spec	Stretch Goal
SIGpara (mas)	3.0	6.0	1.5
SIGpm (mas/yr)	1.0	2.0	0.5
SIGparaRed (mas)	6.0	10.0	3.0

Table 26: The required trigonometric parallax and proper motion accuracy.

# SRD METRICS: AREA AND VISITS



Area/Nvisits design goal:  
18000 sq deg with  
at least 825 visits

In this simulated survey:

An area of 6498 sq deg  
(fO Nvisits:  $0.361 \cdot 18000$ )  
received at least 825 visits

18000 sq deg received at least  
801 visits  
(fO Area:  $0.975 \cdot 825$ )

# SCIENCE METRICS: SRD

- Sky area uniformly covered by the main (WFD) survey
- Median number of visits in all bands across this sky area
- Sky area with nearly uniform visits on short timescales, 40s-30minutes
- Median trigonometric parallax accuracy for  $r=24$  unresolved sources

**Specification:** The sky area uniformly covered by the main survey will include Asky square degrees (Table 22).

Quantity	Design Spec	Minimum Spec	Stretch Goal
Asky ( $\text{deg}^2$ )	18,000	15,000	20,000

Table 22: The sky area uniformly covered by the main survey.

**Specification:** The sum of the median number of visits in each band,  $N_{v1}$ , across the sky area specified in Table 22, will not be smaller than  $N_{v1}$  (Table 23).

Quantity	Design Spec	Minimum Spec	Stretch Goal
$N_{v1}$	825	750	1000

Table 23: The sum of the median number of visits in each band across the sky area specified in Table 22.

**Specification:** At least  $RVA1$  square degrees will have multiple observations separated by nearly uniformly sampled time scales ranging from 40 sec to 30 min (Table 25).

Quantity	Design Spec	Minimum Spec	Stretch Goal
$RVA1$ ( $\text{deg}^2$ )	2,000	1,000	3,000

Table 25: The minimum area with fast (40 sec – 30 min) revisits.

**Specification:** The median trigonometric parallax accuracy across the main survey area for unresolved sources with  $r = 24$  must be at least  $SIG_{para}$  mas. The median proper motion accuracy per coordinate across the main survey area for such sources must be at least  $SIG_{pm}$ . The median trigonometric parallax accuracy across the main survey area for unresolved sources detected only in the  $y$  band (at  $10\sigma$ ) must be at least  $SIG_{paraRed}$  mas. (Table 26).

Quantity	Design Spec	Minimum Spec	Stretch Goal
$SIG_{para}$ (mas)	3.0	6.0	1.5
$SIG_{pm}$ (mas/yr)	1.0	2.0	0.5
$SIG_{paraRed}$ (mas)	6.0	10.0	3.0

Table 26: The required trigonometric parallax and proper motion accuracy.

# SRD METRICS: AREA WITH RAPID REVISITS

# SCIENCE METRICS: SRD

- Sky area uniformly covered by the main (WFD) survey
- Median number of visits in all bands across this sky area
- Sky area with nearly uniform visits on short timescales, 40s-30minutes
- Median trigonometric parallax and proper motion accuracy for  $r=24$  unresolved sources

**Specification:** The sky area uniformly covered by the main survey will include Asky square degrees (Table 22).

Quantity	Design Spec	Minimum Spec	Stretch Goal
Asky (deg <sup>2</sup> )	18,000	15,000	20,000

Table 22: The sky area uniformly covered by the main survey.

**Specification:** The sum of the median number of visits in each band,  $N_{v1}$ , across the sky area specified in Table 22, will not be smaller than  $N_{v1}$  (Table 23).

Quantity	Design Spec	Minimum Spec	Stretch Goal
$N_{v1}$	825	750	1000

Table 23: The sum of the median number of visits in each band across the sky area specified in Table 22.

**Specification:** At least  $RVA1$  square degrees will have multiple observations separated by nearly uniformly sampled time scales ranging from 40 sec to 30 min (Table 25).

Quantity	Design Spec	Minimum Spec	Stretch Goal
$RVA1$ (deg <sup>2</sup> )	2,000	1,000	3,000

Table 25: The minimum area with fast (40 sec – 30 min) revisits.

**Specification:** The median trigonometric parallax accuracy across the main survey area for unresolved sources with  $r = 24$  must be at least  $SIG_{para}$  mas. The median proper motion accuracy per coordinate across the main survey area for such sources must be at least  $SIG_{pm}$ . The median trigonometric parallax accuracy across the main survey area for unresolved sources detected only in the  $y$  band (at  $10\sigma$ ) must be at least  $SIG_{paraRed}$  mas. (Table 26).

Quantity	Design Spec	Minimum Spec	Stretch Goal
$SIG_{para}$ (mas)	3.0	6.0	1.5
$SIG_{pm}$ (mas/yr)	1.0	2.0	0.5
$SIG_{paraRed}$ (mas)	6.0	10.0	3.0

Table 26: The required trigonometric parallax and proper motion accuracy.



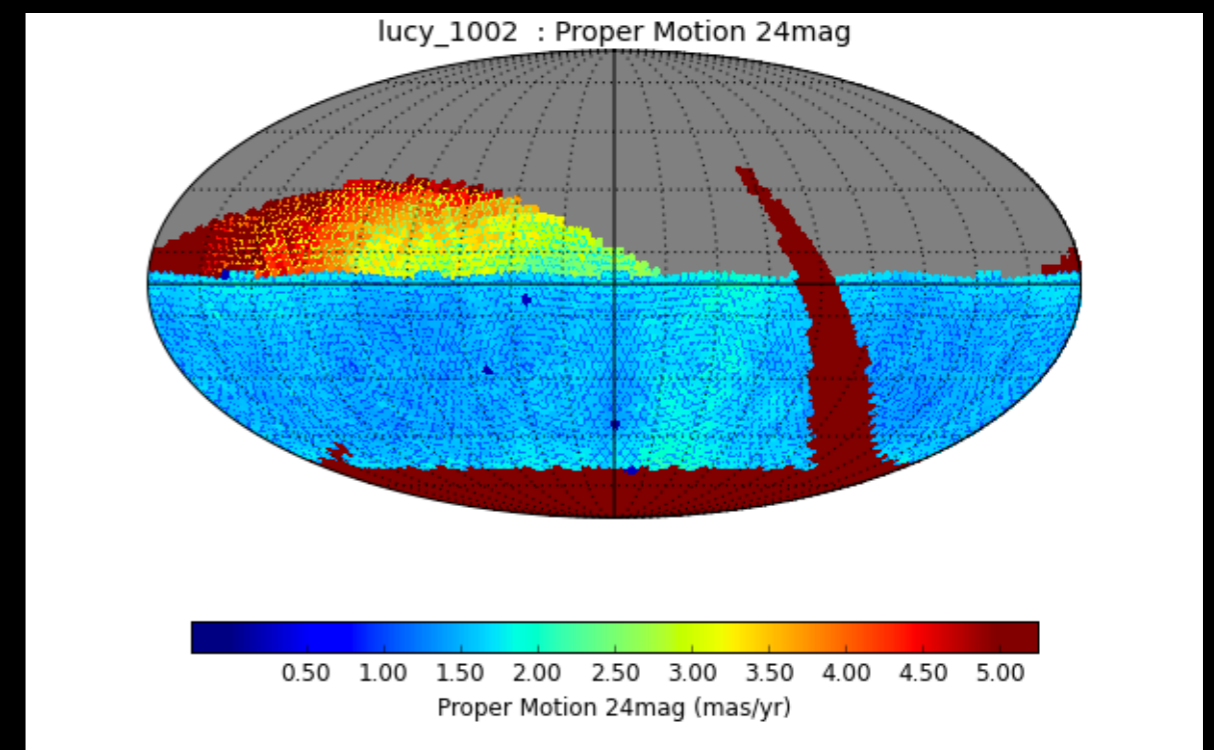
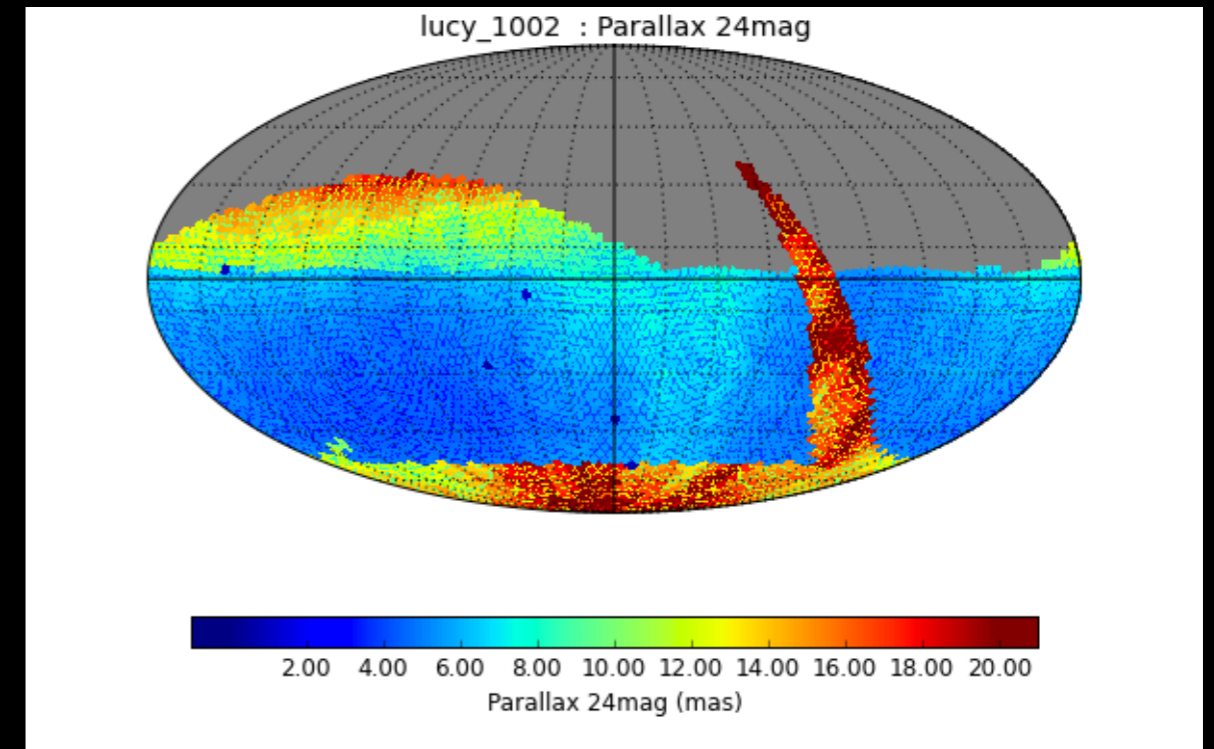
# SRD METRICS: PARALLAX AND PROPER MOTION

Parallax precision estimated from parallax factor (time of observation), astrometric error (SNR) per visit, and errors in resulting fit for parallax.

Proper motion precision estimated from time of visits, astrometric error (SNR) of visit, and errors in resulting fit to a straight line.

(Does not yet account for refraction-induced errors and assumes parallax/pm decoupled)

Design goal:  
3 mas parallax / 1 mas/yr pm



# SRD METRICS: PARALLAX AND PROPER MOTION

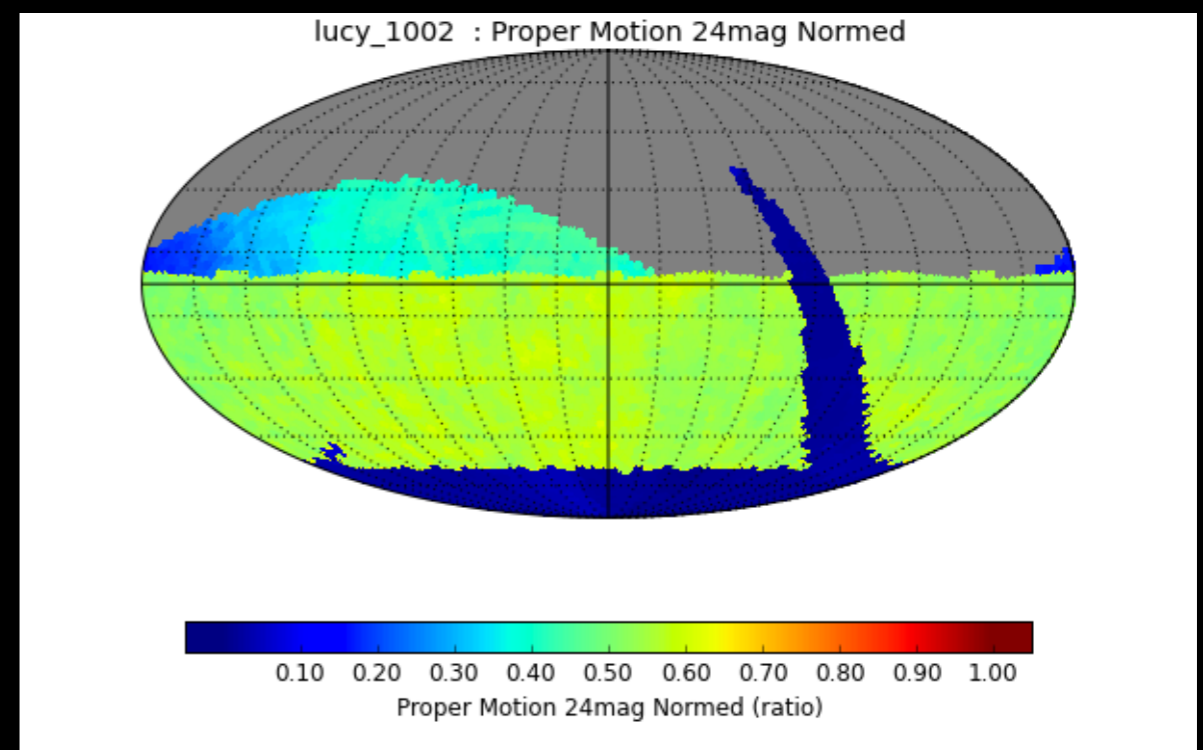
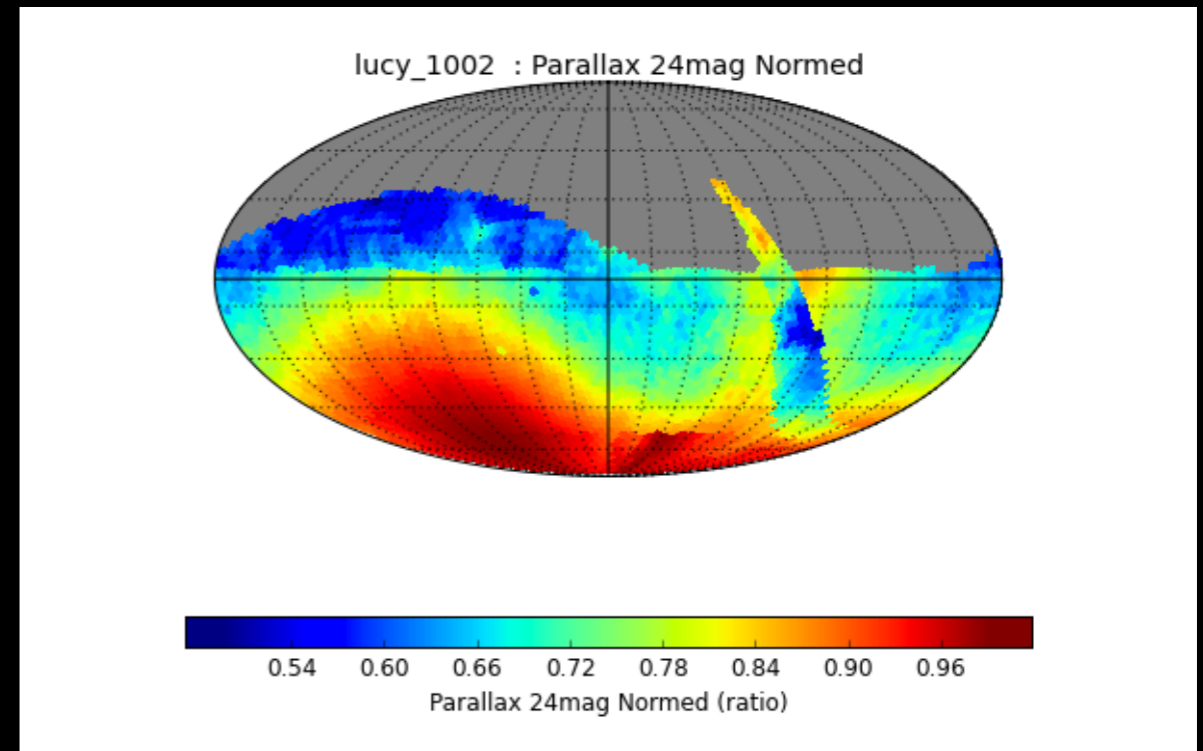
Parallax precision estimated from parallax factor (time of observation), astrometric error (SNR) per visit, and errors in resulting fit for parallax.

Proper motion precision estimated from time of visits, astrometric error (SNR) of visit, and errors in resulting fit to a straight line.

(Does not yet account for refraction-induced errors and assumes parallax/pm decoupled)

Design goal:  
3 mas parallax / 1 mas/yr pm

'Normed' plots : 1 = 'optimal' scheduling



# SCIENCE METRICS: WIDER GOALS

- Beyond these few SRD minimum/design/stretch specifications, additional metrics are desirable to evaluate performance for a wide variety of science goals.
- Without quantitative thresholds, these metrics can be used to rank simulations against each other. Eventually, thresholds could evolve (and will, depending on what is required to enable various science cases).
- Basically infinite possibilities. Can run metrics and use to investigate and rank simulated surveys, but will require scientific input to winnow metrics and determine relative weights before operations.

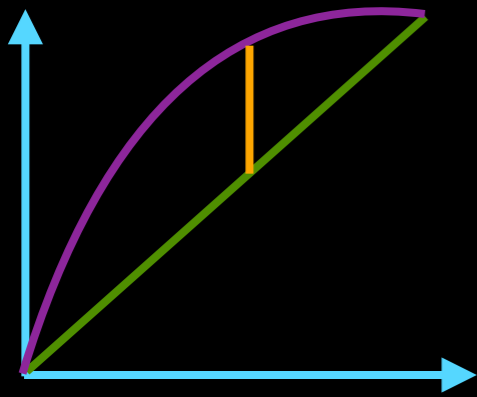


# SCIENCE METRICS: WIDER GOALS DESCRIBED IN THE SRD

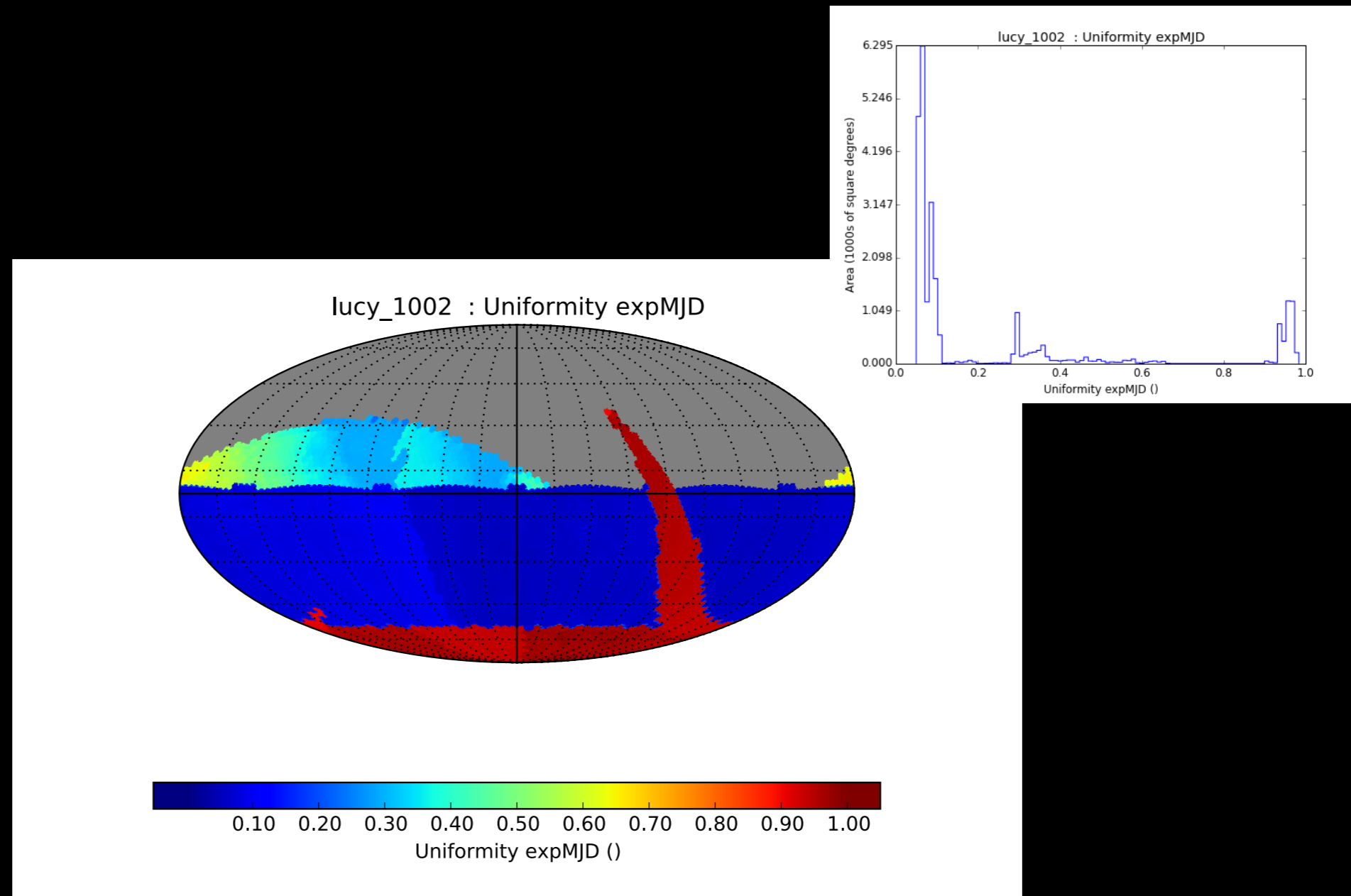
- Uniformity in time and spatial coverage
- Coadded depth in each filter
- Good seeing and low airmass for  $r$  and  $i$  visits
- Number of visits with seeing better than  $X$
- Uniformity and wide range of rotation angles between telescope, camera and sky
- Moving object detection and characterization (visit timing, modified by Solar System Object motion)
- Transient detection and identification (visit timing, coupled with range of filters in those visits)
- Timescales available for color measurements

# SCIENCE METRICS: WIDER GOALS FROM THE SRD

## Temporal uniformity of visits



'Dmax' = max difference  
between cumulative sum of  
visit times vs. straight line  
0 = perfectly uniform  
1 = perfectly nonuniform



# SCIENCE METRICS: WIDER GOALS FROM THE COMMUNITY

- Variable star period recovery
- Phase coverage for variables
- Detection of planetary transits
- Transient detection
- Identification and characterization of time delays in lensed quasars
- Uniformity in Large Scale Structure measurements
- .....

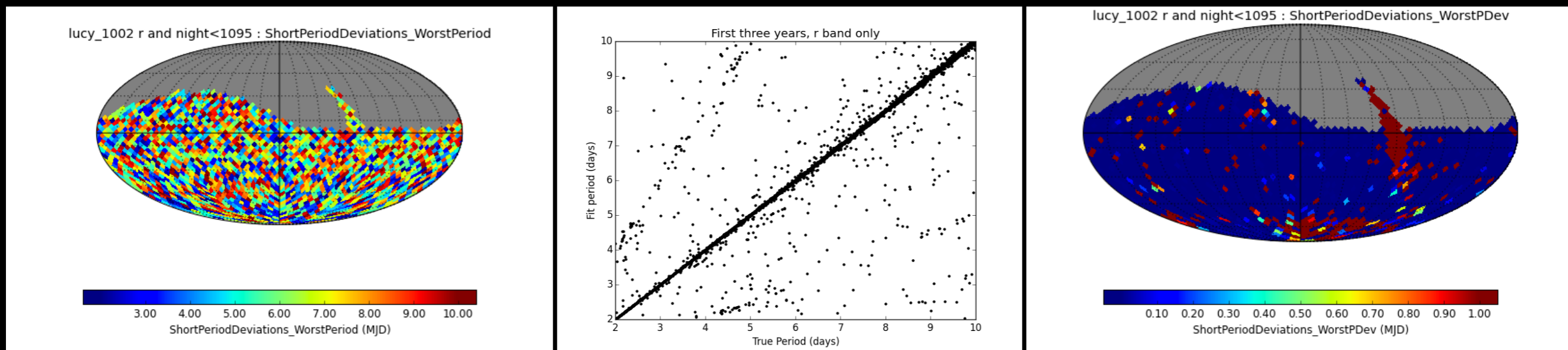
# PERIOD RECOVERY FOR VARIABLES

Generate simple light curve based on a sinusoid (using a range of periods).

Fit period based on observations (with SciPy LS method).

Report PeriodDeviation  $[(\text{fit}-\text{true})/\text{true period}]$ .

Example with periods ranging from 2-10 days, using 3 years of *r* band observations.



# EVALUATING THE SCHEDULER: METRICS

- **Science metrics**

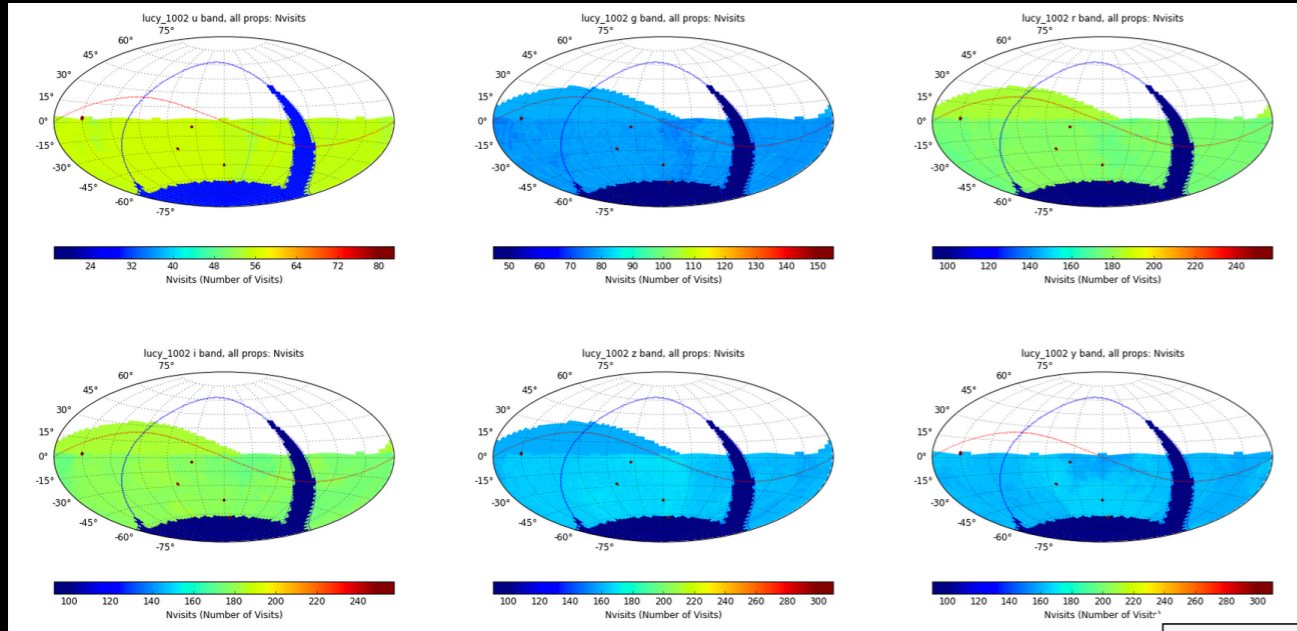
- Evaluate scientific performance of scheduler
- Does survey meet minimum/design/stretch science requirements laid out in Science Requirements Document (SRD)?
- How well does survey perform in a wide variety of other science goals?

- **Technical metrics**

- Evaluate technical performance of scheduler
- Were requested visits obtained? In desired conditions? How efficiently did scheduler obtain visits?

# TECHNICAL METRICS

- Did the scheduler achieve what it was asked to do? How efficiently did was the simulated survey performed?
  - Number of visits per filter per field and 'completeness'
  - Seeing, airmass, sky brightness distributions
  - Hour angle, normalized airmass distributions
  - Slew time distributions
  - Open shutter fraction, number of filter changes



Number of visits  
per field, per filter

Group: C: NVisits; Subgroup: All Props; Slicer: OpsimFieldSlicer

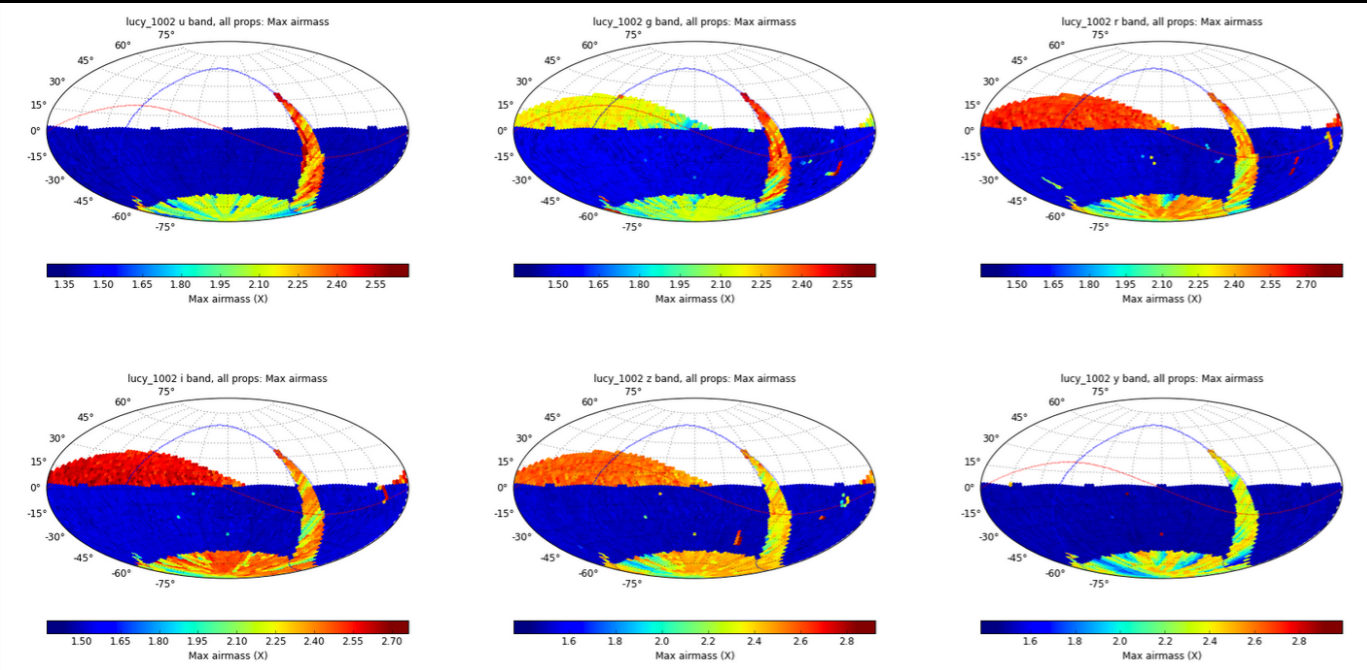
MetricName	Metadata	Median	Mean	Rms	m3Sigma	p3Sigma	Count	25th%ile	75th%ile	Min	Max
Nvisits	u band, all props	55.00	67.26	393.57	0	5	2817	55.00	56.00	30.00	10031.00
Nvisits	g band, all props	78.00	76.16	141.86	0	5	3338	77.00	79.00	30.00	4146.00
Nvisits	r band, all props	176.00	165.32	285.57	0	5	3339	174.00	178.00	30.00	8247.00
Nvisits	i band, all props	178.00	166.12	277.68	0	5	3339	175.00	180.00	30.00	7940.00
Nvisits	z band, all props	164.00	156.47	347.63	0	5	3339	160.00	167.00	30.00	9835.00
Nvisits	y band, all props	162.00	149.65	281.03	0	5	2817	160.00	164.00	30.00	7199.00

Completeness

Group: F: Completeness; Subgroup: WFD; Slicer: OpsimFieldSlicer

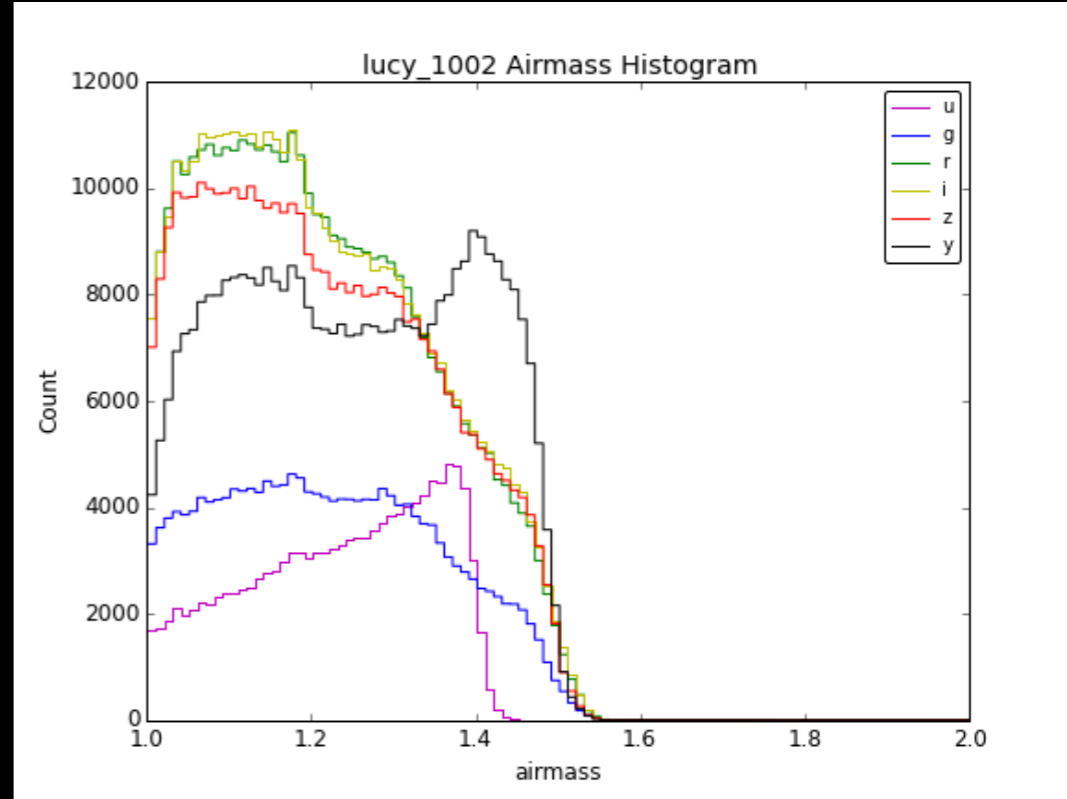
	Completeness_u	Completeness_g	Completeness_r	Completeness_i	Completeness_z	Completeness_y	Completeness_Joint
	WFD only	WFD only	WFD only	WFD only	WFD only	WFD only	WFD only
TableFraction 0 == P	0	0	0	0	0	0	0
TableFraction 0 < P < 0.1	0	0	0	0	0	0	0
TableFraction 0.1 <= P < 0.2	0	0	0	0	0	0	0
TableFraction 0.2 <= P < 0.3	0	0	0	0	0	0	0
TableFraction 0.3 <= P < 0.4	0	0	0	0	0	0	0
TableFraction 0.4 <= P < 0.5	0	0	0	0	0	0	0
TableFraction 0.5 <= P < 0.6	0	0	0	0	0	0	0
TableFraction 0.6 <= P < 0.7	0	0	0	0	0	0	0
TableFraction 0.7 <= P < 0.8	0	0	0	0	0	0	0
TableFraction 0.8 <= P < 0.9	0	0	0	0	0	0	0
TableFraction 0.9 <= P < 1	1173	2283	2290	2287	53	159	2290
TableFraction 1 == P	1117	7	0	3	91	319	0
TableFraction 1 < P	3	3	3	3	2149	1815	3



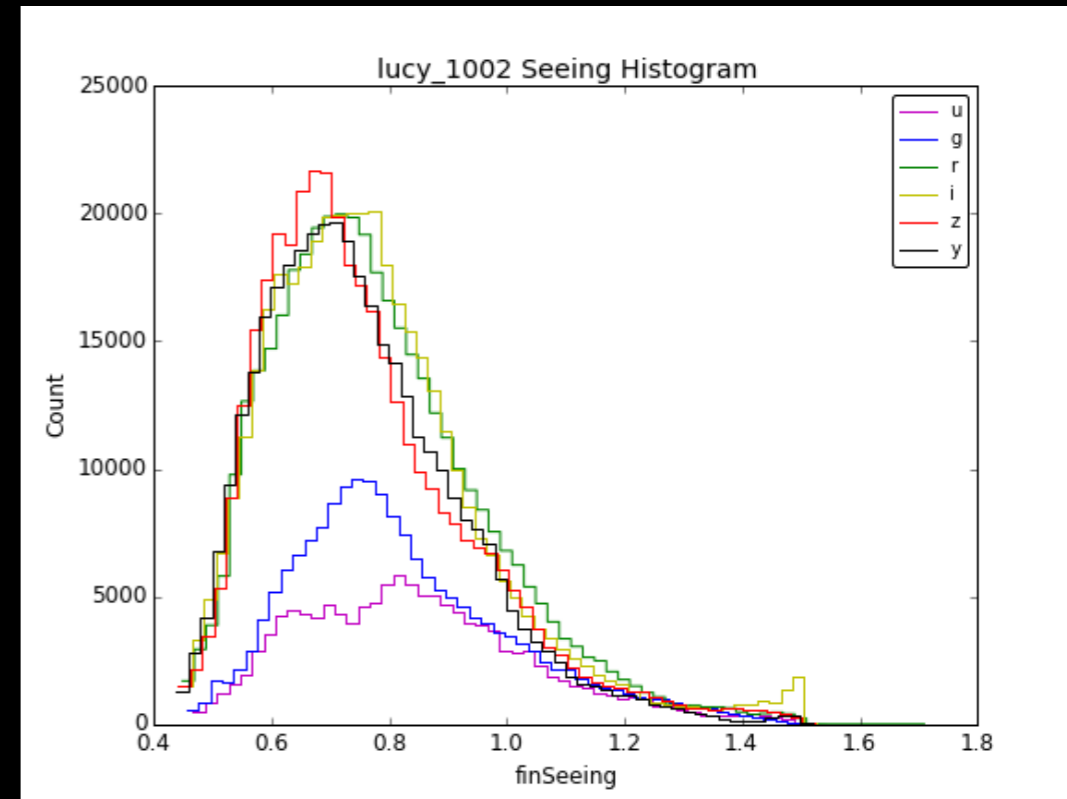
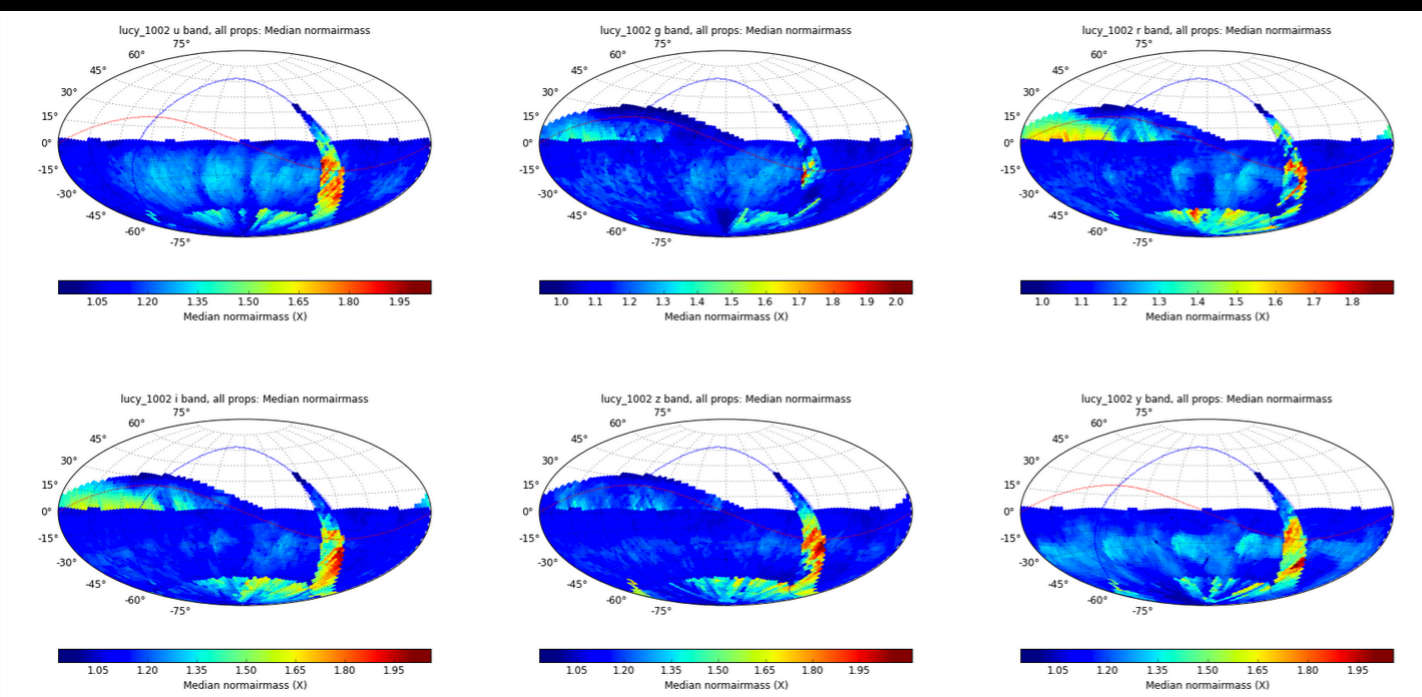


Max airmass per field (and mean, min..).

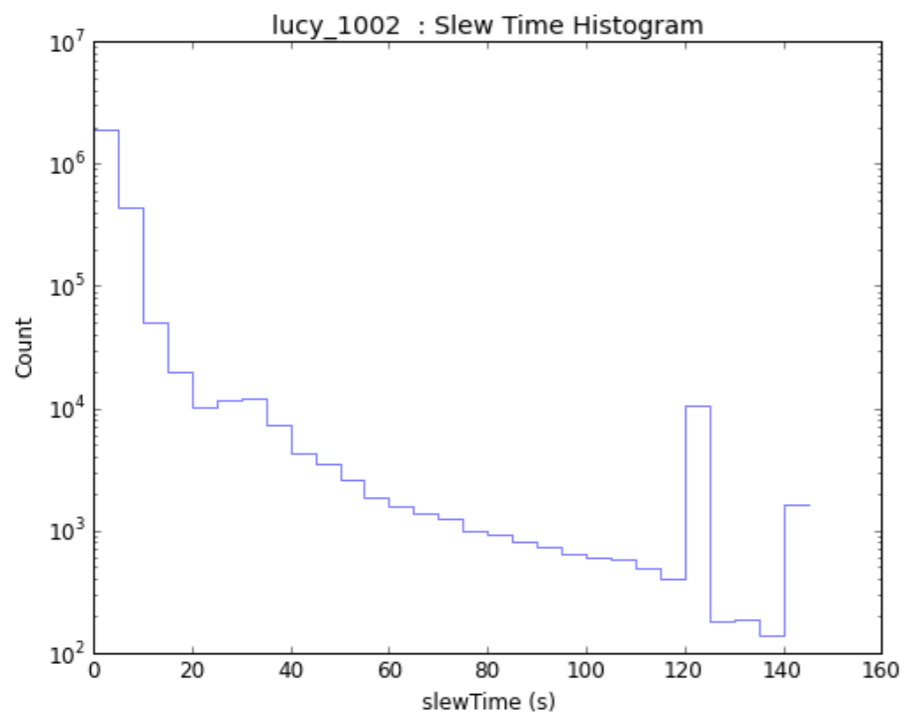
Median normalized airmass (airmass/  
minimum airmass possible for each field).



Airmass (above) and seeing (below) distributions for all visits



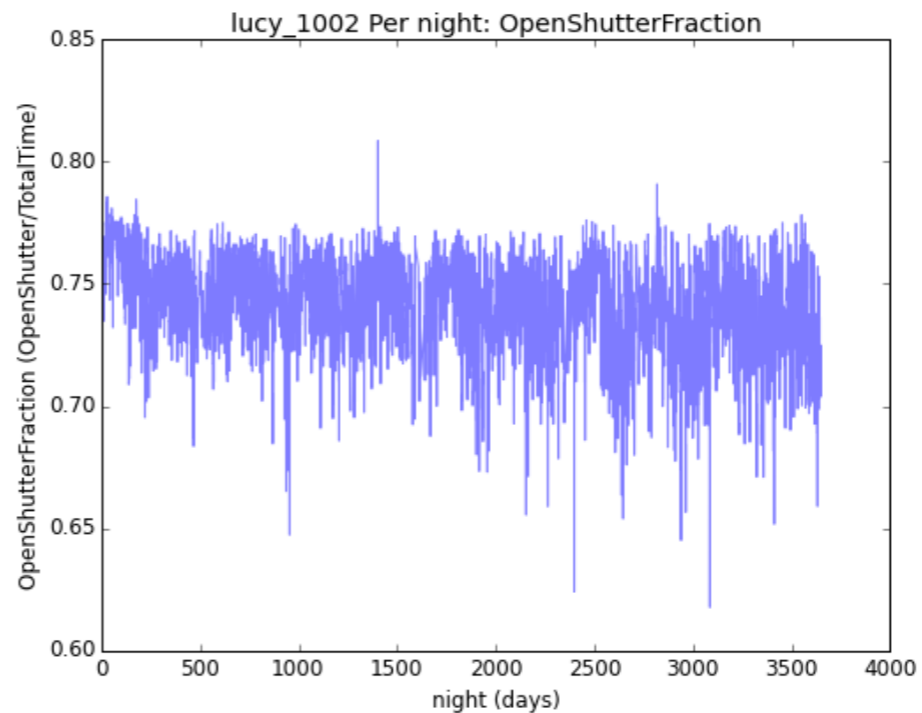
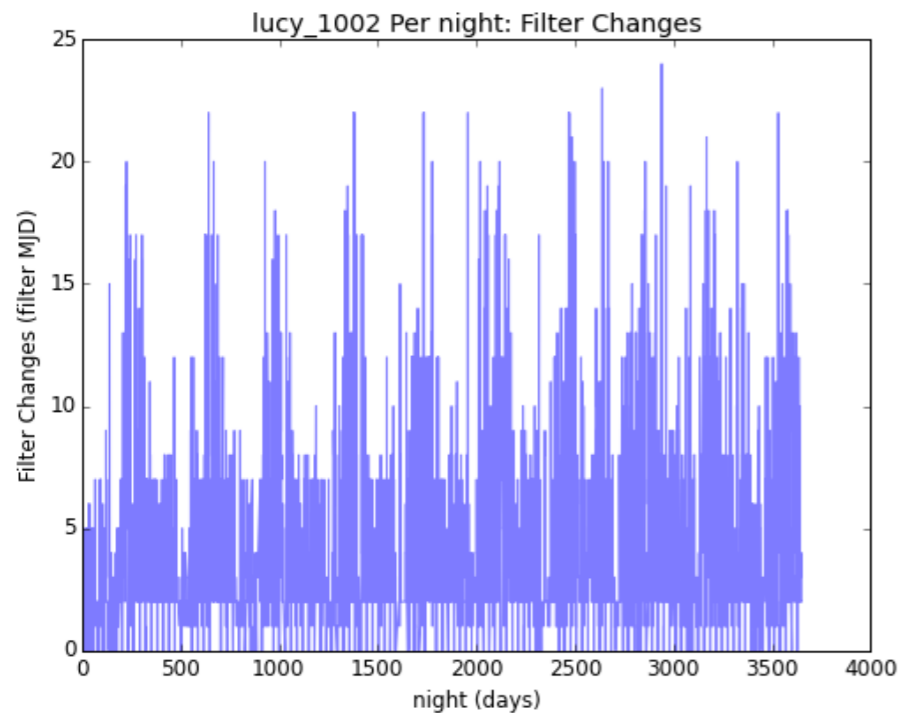




Group: *L: Slew*; Subgroup: *Slew Activity*; Slicer: *UniSlicer*

	ActivePerc	ActiveAve	Max	ActivePerc in crit	ActiveAve in crit	Total Ave	Contribution
<b>DomAlt</b>	91.14	3.41	60.24	2.19	7.89	3.11	0.21
<b>DomAz</b>	91.37	5.36	121.96	30.08	8.46	4.90	3.30
<b>TelAlt</b>	91.14	1.71	30.12	49.95	1.78	1.55	0.24
<b>TelAz</b>	91.14	1.63	52.32	8.59	1.60	1.48	0.03
<b>Rotator</b>	0.50	11.54	29.21	0.02618	0.02618	0.06	0.0000
<b>Filter</b>	0.48	120.00	120.00	0.48	120.00	0.58	10.59
<b>TelOpticsOL</b>	91.14	0.80	29.12	0.0065	5.99	0.73	0.0004
<b>Readout</b>	100.00	2.00	2.00	6.97	2.00	2.00	0.04
<b>Settle</b>	92.90	3.00	3.00	60.27	3.00	2.79	0.83
<b>TelOpticsCL</b>	1.24	20.00	20.00	1.24	20.00	0.25	0.76

Slew time histogram (note big bump for filter changes).  
 Number of filter changes per night.  
 Open shutter fraction per night.



# EVALUATING THE SCHEDULER: METRICS

- **Science metrics**

- Evaluate scientific performance of scheduler
- Does survey meet minimum/design/stretch science requirements laid out in Science Requirements Document (SRD)?
- How well does survey perform in a wide variety of other science goals?

- **Technical metrics**

- Evaluate technical performance of scheduler
- Were requested visits obtained? In desired conditions? How efficiently did scheduler obtain visits?

- Are we thinking of the right things for metrics? Are our metrics useful? Are we missing metrics?
- Currently we're evaluating metrics at the end of the simulated survey to compare/rank/update observing strategy for the next simulation. How do we (should we) move toward folding metrics into scheduler as a feedback mechanism?