

Simulator Output Format & Content

Please dont change the following page without talking to Srini Chandrasekharan. This page is being used as a specifications document

SPECIFICATIONS for Database Registration process

Workflow Considerations

We will need to consider how database registration (or creation of a product from the Simulator DB that is usable by MAF) will fit into the overall workflow.

1. Do we need to create a separate "pre-processing" codebase? Should this be incorporated by MAF as part of "processing", or should it be incorporated by the Simulator as "post_processing" to the simulation?
2. If we keep the "summary_" table in the database will the database become too large at some point? Do we need to start keeping simulated surveys in separate databases?
3. How does OpSim best transition from "output_" and SSTAR to "summary_" and MAF without inhibiting productivity (holding up current analysis).

Audiences / Use Cases

1. OpSim
 - a. SSTAR performs "processing" that produces an extra table ("output_", to evolve to "summary_") that is written to a file as well as single-session files that are portable.
 - b. SSTAR acts on the "output_" table as well as the database as a whole, keeping the "output_" table as an extra table in the database, to generate its output.
 - c. If OpSim is to phase out SSTAR, then a new tool needs to be implemented to generate the "summary_" table (and other exported files as specified in Deliverable below) and MAF will have to completely replicate SSTAR.
2. MAF Team
 - a. The starting input is an SQLite file of the "output_" table (now "summary_"), assuming pre-processing is already completed.
3. Science Community user
 - a. To date the expectation is that the user will download a "summary_" file from the OpSim Data Access website, install MAF and do analysis.
 - b. Now that the Simulator is open source, if the user creates his/her own simulations, they will also need this pre-processing code before they can use MAF

Specifications

- A new table will be created that is called "summary_*machinename_sessionID*"
- Columns will be copies of those from the ObsHistory table along with propConf from Proposal that corresponds with propID
- Specific columns:

Field in "output_"	Field in "ObsHistory"	Proposed Fields in "summary_"	Notes	
obsHistID	obsHistID	obsHistID		
sessionID	Session_sessionID	sessionID		
filter	filter	filter		
expDate	expDate	expDate		
night	night	night		
visitTime	visitTime	visitTime	This should be 34sec but is zero	
visitExpTime	visitExpTime	visitExpTime	This should be 30 sec but is 34 sec	
finRank	finRank	finRank		
finSeeing	finSeeing	finSeeing		
transparency	transparency	transparency		
airmass	airmass	airmass		
vSkyBright	vSkyBright	vSkyBright		
filtSkyBright	filtSkyBright	filtSkyBrightness		
rotSkyPos	rotSkyPos	rotSkyPos		
lst	lst	lst		

altitude	alt	altitude		
azimuth	az	azimuth		
dist2Moon	dist2Moon	dist2Moon		
solarElong	solarElong	solarElong		
moonRA	moonRA	moonRA		
moonDec	moonDec	moonDec		
moonAlt	moonAlt	moonAlt		
moonAZ	moonAZ	moonAZ		
moonPhase	moonPhase	moonPhase	actually the illumination; maybe call moonIllumination - this would mean a change to the Simulator code's use of this field so for the Summary table it will stay the same and get better documentation	
sunAlt	sunAlt	sunAlt		
sunAz	sunAZ	sunAZ		
phaseAngle	phaseAngle	phaseAngle		
rScatter	rScatter	rScatter		
mieScatter	mieScatter	mieScatter		
moonIllum	moonIllum	moonIllum	rename to a parameter - also means a change to OpSim so better documentation is needed	
moonBright	moonBright	moonBright		
darkBright	darkBright	darkBright		
rawSeeing	rawSeeing	rawSeeing		
wind	wind	wind		
humidity	humidity	humidity		
fieldID	Field_fieldID	fieldID	key to Field table	
fieldRA		fieldRA	from Field table	
fieldDec		fieldDec	from Field table	
propID		propID	keyed by ObsHistID to ObsHistory_Proposal table	
			omit name of proposal config file from Proposal table	
expMJD	expMJD	expMJD		
slewDist		slewDist	from SlewHistory table	
slewTime		slewTime	from SlewHistory table	
5sigma		fiveSigmaDepth	calculated (will be used in the future)	move to ObsHistory
perry_skybrightness			calculated (will be used in the future)	do in MAF
5sigma_ps			calculated (will be used in the future)	do in MAF
skybrightness_modified			calculated (will be used in the future)	do in MAF
5sigma_modified			calculated (will be used in the future)	do in MAF
hexdithra		ditheredRA	calculated (when we have dithering)	move to ObsHistory
hexdithdec		ditheredDec	calculated (when we have dithering)	move to ObsHistory
vertex			calculated (when we have dithering)	remove

- Files to be created
 - ASCII dat format of the "summary_" table (already implemented in SSTAR; needs changes above)
 - MySQL .sql format of all OpsimDB tables including "summary_" table for a sessionID (already implemented in SSTAR; needs changes above)
 - MySQL .sql format of the "summary_" table (already implemented in SSTAR; needs changes above)
 - SQLite format file containing all tables including "summary_" table

MEETING ON APRIL 10, 2014

Attendees: Peter, Simon, Kem, Cathy, Lynne, Srini, Steve

1. Deliverable of Opsim - What should it be?
 - a. Both output table & the entire DB in sqlite format + output table in dat format
 - b. csv discussion - Deferred
2. What happens to the additional columns ?
 - a. dithra & dithdec etc will be added to preprocessing
 - b. extra columns will goto preprocessing

Deliverable

Preprocessing layer which will take as input the OpSim DB and convert that to:

- Flat dat (output_ table) [already implemented]
- MySQL .sql (one for all tables including output tables for a sessionID, and one for the output table) - this is already being done phase out
- Add sqlite file containing (all tables including output_ table)

NOTES BEFORE THE MEETING

Meeting Objective

Discuss needs and wants associated with producing content (output) from the Simulator that is usable by both the team and the community, and that will be accessed by MAF.

Background

1. Generating "extra" columns
 - a. 5sigma, perry_skybrightness, 5sigma_ps, skybrightness_modified, 5sigma_modified, hexdithra, hexdithdec, vertex
 - b. these quantities are not used or required by the Simulator, and philosophically no post-processing should add/change the OpsimDB database, so these are created by gen_output.py in SSTAR.
 - c. the "output_machine_id" table is a copy of the ObsHistory table plus the fields in (a)
2. Sky Brightness models & fields
 - a. when a unified model is adopted by the project, it will be added to the Simulator and those fields will be recorded in the ObsHistory table.
 - b. the different ways of computing sky brightness have different advantages - skybrightness_modified and 5sigma_modified were developed in an effort represent the "best" value as a combination of the other two. Here is the description from the SSTAR documentation:

The Operations Simulator uses two methods of calculating the sky brightness at the time of a visit: OpSimSky and ETCSky.

Before each visit, the sky brightness in the V filter, VskyBright, is evaluated using the Johnson V band calculated from the Krisciunas and Schaeffer model, with a few modifications. This model uses the Moon phase, angular distance between the field and the Moon and the field's airmass to calculate added brightness to the zero-Moon, zenith sky brightness (e.g. Krisciunas 1997, PASP, 209, 1181; Krisciunas and Schaefer 1991, PASP, 103, 1033; Benn and Ellison 1998, La Palma Technical Note 115).

From VskyBright there are two different methods to evaluate the sky brightness in a particular band. One method is to take measurements of the color of the sky as a function of lunar phase, and use these to correct VskyBright to a particular filter. This is the approach used for the OpSim sky brightness, OpSimSky.

An alternate method, ETCSky, is to use sky brightness measurements taken at various phases of the moon in many filters and calculate empirical fits to the sky brightness in each band. This method is what has been used to create the LSST Exposure Time Calculator (ETC) and has the advantage of accounting for the fact that the night sky does not behave similarly in all bands by including the angular dependence as a function of filter bandpass. The plots and tables in this document report the sky brightness from the ETCSky method.

Questions

1. Who are we creating the "output" for? should OpSim/MAF be accessing the DB directly? what about community users?
2. What information is needed by OpSim (all tables), and the community (just ObsHistory?)
3. Do we need one "materialized" table for simplicity or can all tables be distributed in one file?
4. What does SSTAR use? (not output*)
5. While extensibility and inter-backend-operability is key for MAF what would be the ideal input format?

6. How do you include proposal information? Are duplicate rows acceptable?