Aggregating MAF results

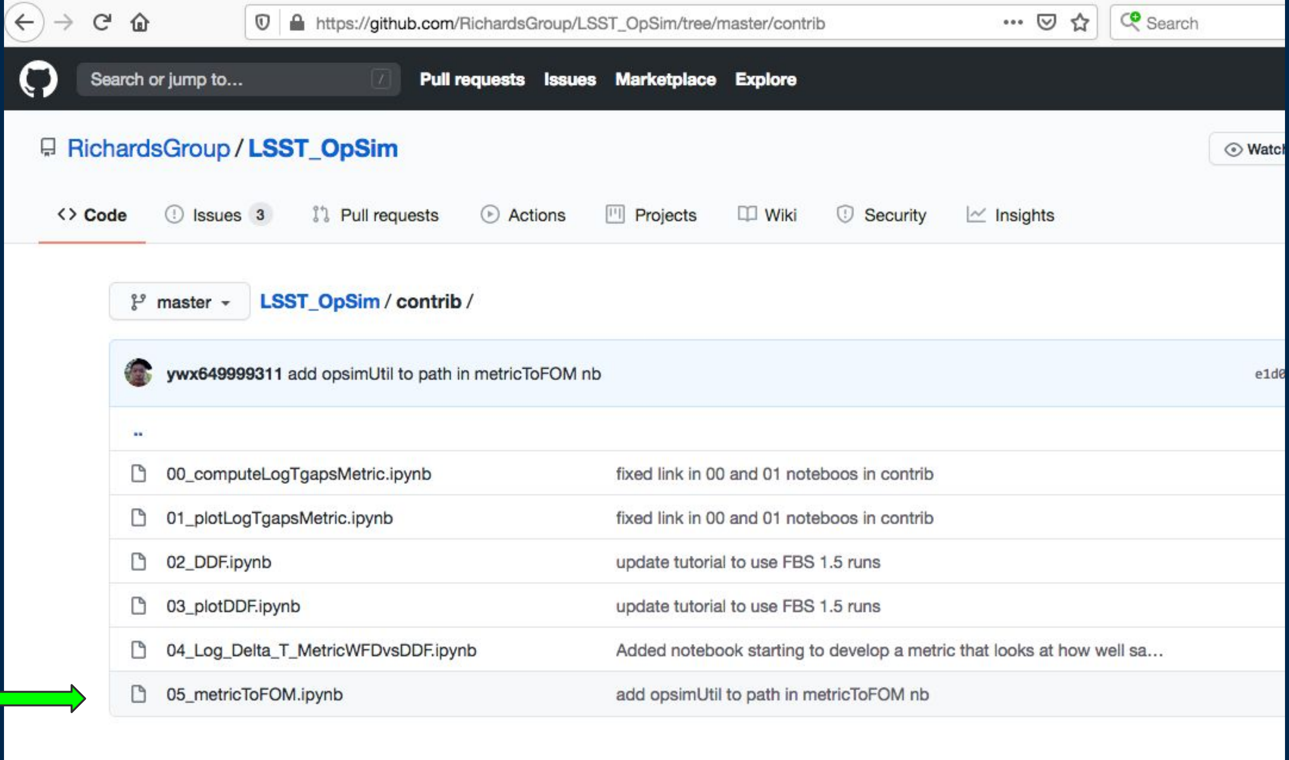
Will Clarkson (University of Michigan-Dearborn)

Fabio Ragosta (INAF and University of Naples “Federico II”)

Xiaolong Li (University of Delaware)

Federica Bianco (University of Delaware)

https://github.com/RichardsGroup/LSST_OpSim/blob/master/contrib/05_metricToFOM.ipynb



https://github.com/RichardsGroup/LSST_OpSim/tree/master/contrib

RichardsGroup / LSST_OpSim

LSST_OpSim / contrib /

ywx649999311 add opsimUtil to path in metricToFOM nb e1d0

..

00_computeLogTgapsMetric.ipynb	fixed link in 00 and 01 noteboos in contrib
01_plotLogTgapsMetric.ipynb	fixed link in 00 and 01 noteboos in contrib
02_DDF.ipynb	update tutorial to use FBS 1.5 runs
03_plotDDF.ipynb	update tutorial to use FBS 1.5 runs
04_Log_Delta_T_MetricWFDvsDDF.ipynb	Added notebook starting to develop a metric that looks at how well sa...
05_metricToFOM.ipynb	add opsimUtil to path in metricToFOM nb

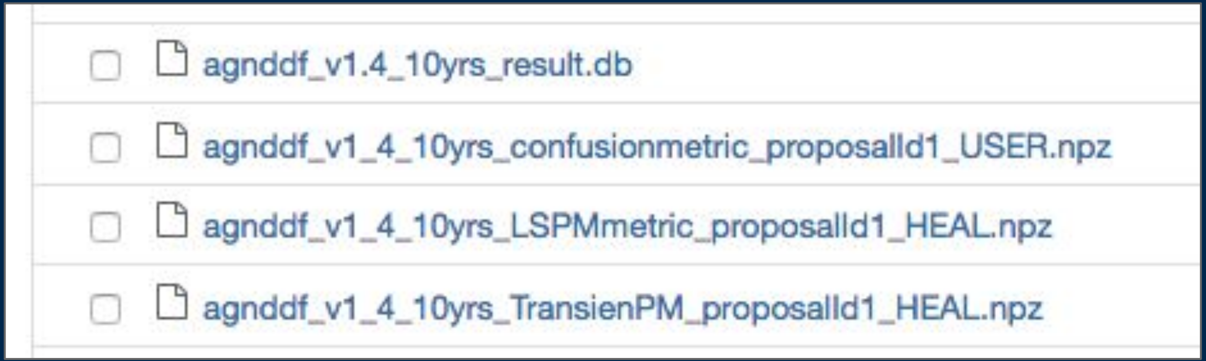
https://github.com/RichardsGroup/LSST_OpSim/blob/master/contrib/05_metricToFOM.ipynb

Use case: you have a big pile of evaluated metrics to be combined into your science figures of merit. You now want to summarize the runs for side-by-side comparison, while preserving the relative performance for each metric.

I will demo an example method* to accomplish this.

*Credit: Fabo Ragosta
and Xiaolong Li

In this case, three metrics have already been run on 76 OpSim runs → one .db and three .npz files per OpSim:

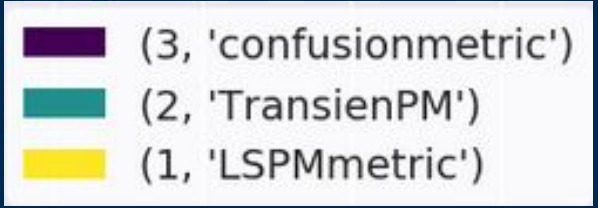


□ agnddf_v1.4_10yrs_result.db

□ agnddf_v1_4_10yrs_confusionmetric_proposalId1_USER.npz

□ agnddf_v1_4_10yrs_LSPMmetric_proposalId1_HEAL.npz

□ agnddf_v1_4_10yrs_TransienPM_proposalId1_HEAL.npz

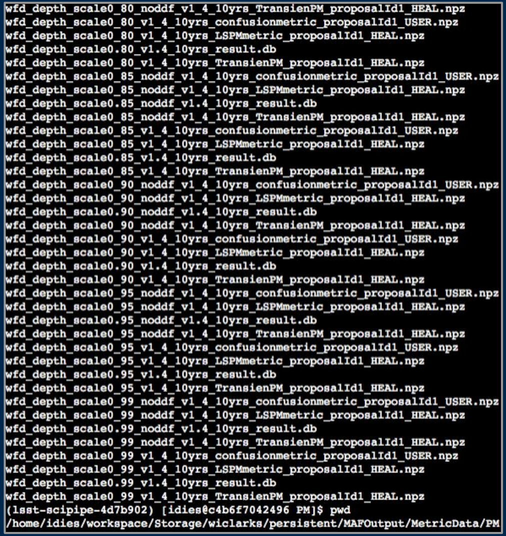


🟪 (3, 'confusionmetric')

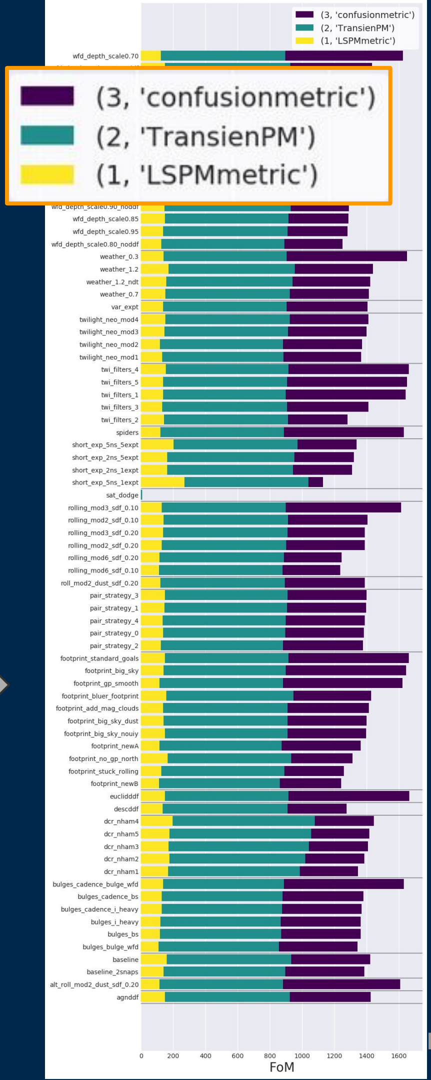
🟩 (2, 'TransienPM')

🟨 (1, 'LSPMmetric')

metricToFOM: Aggregates all MAF results (*.npz, *.db) into one FoM per metric per OpSim and plots a summary figure. Optionally outputs FoMs to disk.



```
wfd_depth_scale0_80_noddf_v1_4_10yrs_TransienPM_proposalId1_HEAL.npz
wfd_depth_scale0_80_v1_4_10yrs_confusionmetric_proposalId1_USER.npz
wfd_depth_scale0_80_v1_4_10yrs_LSPMmetric_proposalId1_HEAL.npz
wfd_depth_scale0.80_v1.4_10yrs_result.db
wfd_depth_scale0_80_v1_4_10yrs_TransienPM_proposalId1_HEAL.npz
wfd_depth_scale0_85_noddf_v1_4_10yrs_confusionmetric_proposalId1_USER.npz
wfd_depth_scale0_85_noddf_v1_4_10yrs_LSPMmetric_proposalId1_HEAL.npz
wfd_depth_scale0.85_noddf_v1.4_10yrs_result.db
wfd_depth_scale0_85_noddf_v1_4_10yrs_TransienPM_proposalId1_HEAL.npz
wfd_depth_scale0_85_v1_4_10yrs_confusionmetric_proposalId1_USER.npz
wfd_depth_scale0_85_v1_4_10yrs_LSPMmetric_proposalId1_HEAL.npz
wfd_depth_scale0.85_v1.4_10yrs_result.db
wfd_depth_scale0_85_v1_4_10yrs_TransienPM_proposalId1_HEAL.npz
wfd_depth_scale0_90_noddf_v1_4_10yrs_confusionmetric_proposalId1_USER.npz
wfd_depth_scale0_90_noddf_v1_4_10yrs_LSPMmetric_proposalId1_HEAL.npz
wfd_depth_scale0.90_noddf_v1.4_10yrs_result.db
wfd_depth_scale0_90_noddf_v1_4_10yrs_TransienPM_proposalId1_HEAL.npz
wfd_depth_scale0_90_v1_4_10yrs_confusionmetric_proposalId1_USER.npz
wfd_depth_scale0_90_v1_4_10yrs_LSPMmetric_proposalId1_HEAL.npz
wfd_depth_scale0.90_v1.4_10yrs_result.db
wfd_depth_scale0_90_v1_4_10yrs_TransienPM_proposalId1_HEAL.npz
wfd_depth_scale0_95_noddf_v1_4_10yrs_confusionmetric_proposalId1_USER.npz
wfd_depth_scale0_95_noddf_v1_4_10yrs_LSPMmetric_proposalId1_HEAL.npz
wfd_depth_scale0.95_noddf_v1.4_10yrs_result.db
wfd_depth_scale0_95_noddf_v1_4_10yrs_TransienPM_proposalId1_HEAL.npz
wfd_depth_scale0_95_v1_4_10yrs_confusionmetric_proposalId1_USER.npz
wfd_depth_scale0_95_v1_4_10yrs_LSPMmetric_proposalId1_HEAL.npz
wfd_depth_scale0.95_v1.4_10yrs_result.db
wfd_depth_scale0_95_v1_4_10yrs_TransienPM_proposalId1_HEAL.npz
wfd_depth_scale0_99_noddf_v1_4_10yrs_confusionmetric_proposalId1_USER.npz
wfd_depth_scale0_99_noddf_v1_4_10yrs_LSPMmetric_proposalId1_HEAL.npz
wfd_depth_scale0.99_noddf_v1.4_10yrs_result.db
wfd_depth_scale0_99_noddf_v1_4_10yrs_TransienPM_proposalId1_HEAL.npz
wfd_depth_scale0_99_v1_4_10yrs_confusionmetric_proposalId1_USER.npz
wfd_depth_scale0_99_v1_4_10yrs_LSPMmetric_proposalId1_HEAL.npz
wfd_depth_scale0.99_v1.4_10yrs_result.db
wfd_depth_scale0_99_v1_4_10yrs_TransienPM_proposalId1_HEAL.npz
(lsst-scipipe-4d7b902) [idies@c4b6f7042496 PM]$ pwd
/home/idies/workspace/Storage/wiclarks/persistent/MAFOutput/MetricData/PM
```



wfd_

weather_

twilight_

twi_

short_

rolling_

pair_

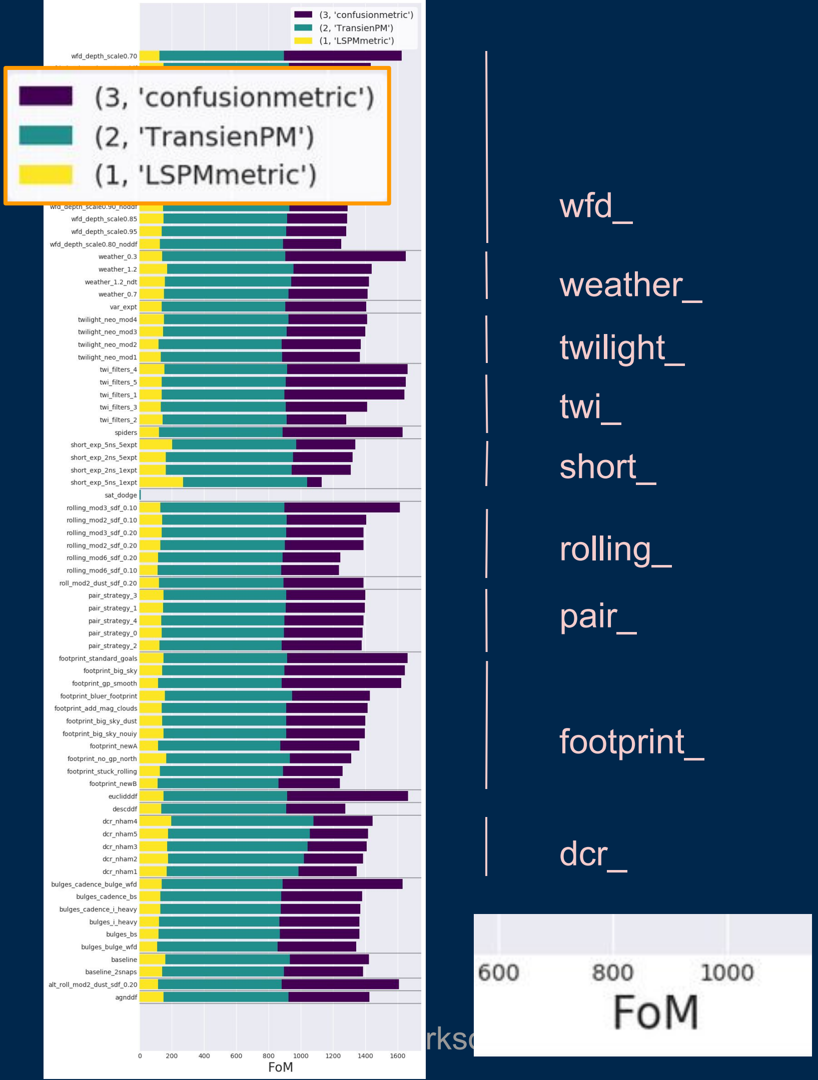
footprint_

dcr_

metricToFOM: Aggregates all MAF results (*.npz, *.db) into one FoM per metric per OpSim and plots a summary figure. Optionally outputs FoMs to disk.

$$\mathrm{FoM}_k = N_k \sum_{i}^{\mathrm{healpix}} (\mathrm{metric}_k)_i$$

- By default, grouped by OpSim “family”
- By default, ranked within each family
- Metrics are summed over the fields, optionally renormalized, producing one FoM_k per metric
- Lower and upper thresholds can be specified for each metric
- Choices of renormalization (default: N_k=1)

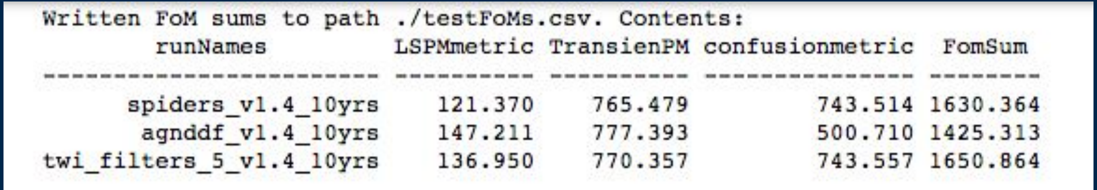


Metrics Hackathon, 2020-08-06: Aggregating MAF results

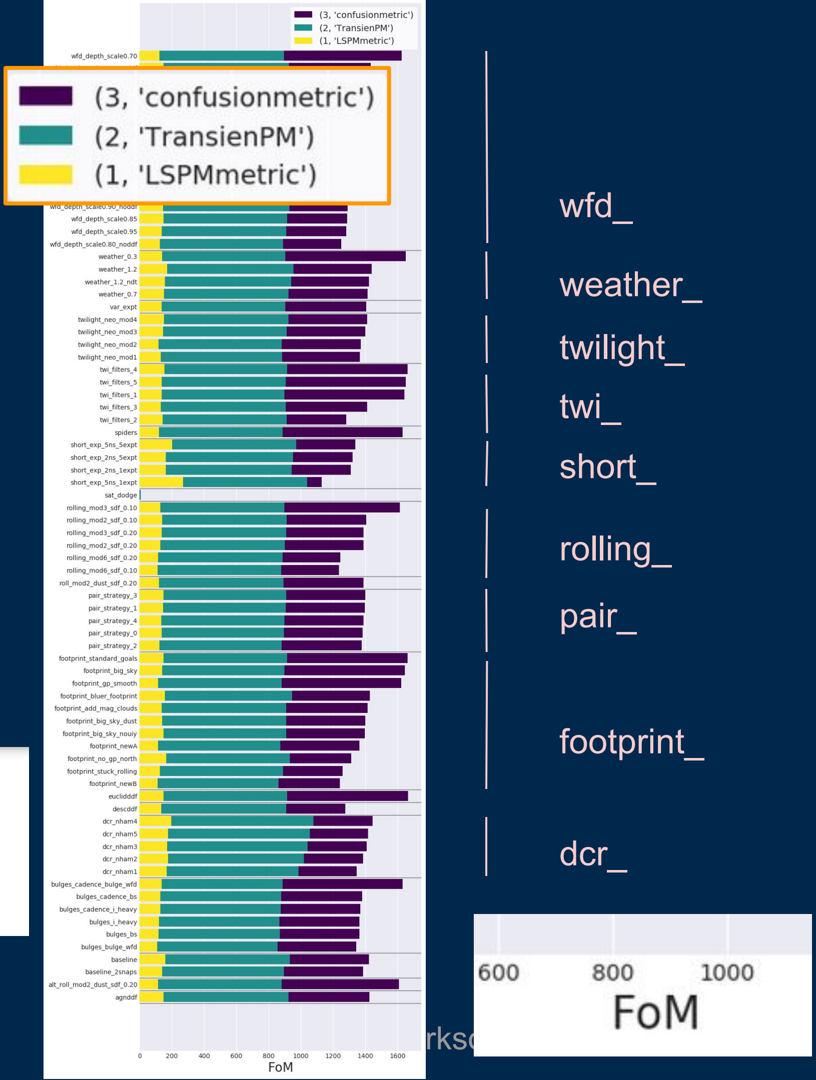
metricToFOM: Aggregates all MAF results (*.npz, *.db) into one FoM per metric per OpSim and plots a summary figure. Optionally outputs FoMs to disk.

$$\mathrm{FoM}_k = N_k \sum_i^{\mathrm{healpix}} (\mathrm{metric}_k)_i$$

- The FoM set can be output to disk, along with its sum over all the metrics, one row per OpSim. Example output:



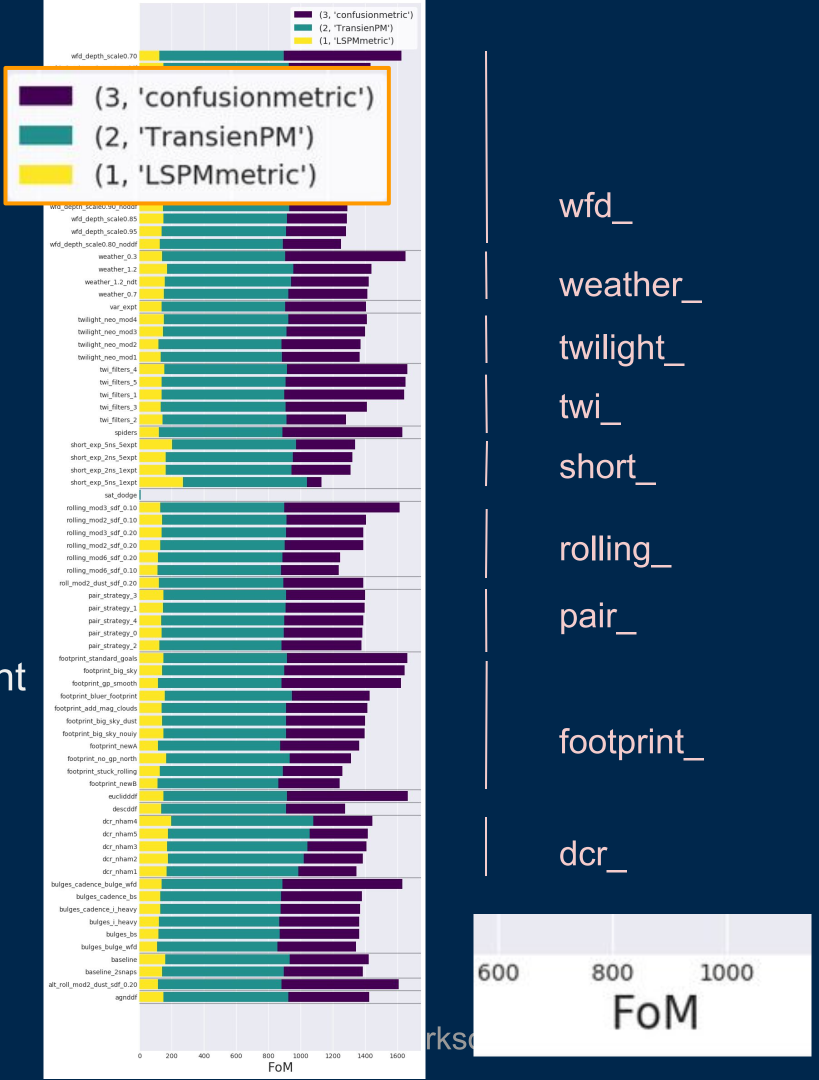
```
Written FoM sums to path ./testFoMs.csv. Contents:
         runNames          LSPMmetric TransienPM confusionmetric  FomSum
------------------------- ---------- ---------- --------------- ---------
     spiders_v1.4_10yrs     121.370    765.479          743.514 1630.364
      agnddf_v1.4_10yrs     147.211    777.393          500.710 1425.313
twi_filters_5_v1.4_10yrs    136.950    770.357          743.557 1650.864
```



metricToFOM: Aggregates all MAF results (*.npz, *.db) into one FoM per metric per OpSim and plots a summary figure. Optionally outputs FoMs to disk.

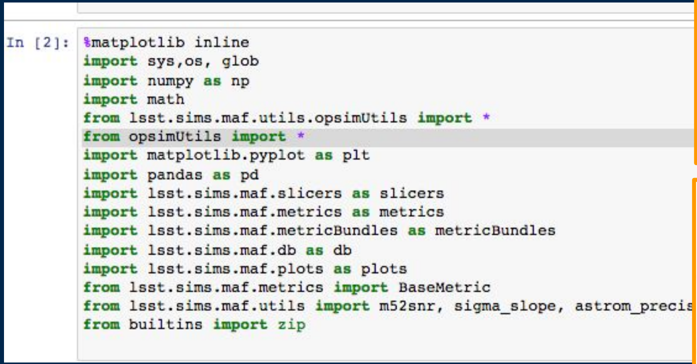
$$\mathrm{FoM}_k = N_k \sum_{i}^{\text{healpix}} (\mathrm{metric}_k)_i$$

- Default search path: /home/idies/workspace/Storage/{username}/persistent /MAFOutput/MetricData/PM/'
- BOTH *.npz and *.nb MAF outputs are needed



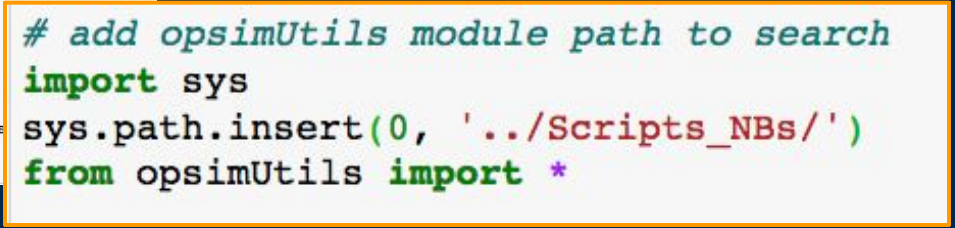
Learning from experience: importing the correct opsimUtils

- This is opsimUtils in the LSST_OpSim repository (by Gordon and Weixiang)
  - (it is NOT the opsimUtils in lsst.sims.maf.utils, as I had first thought !!)
- I recommend copying (or symlinking) the file below into the working directory: /path/to/LSST_OpSim/Scripts_NBs/opsimUtils.py



```
In [2]: %matplotlib inline
        import sys,os, glob
        import numpy as np
        import math
        from lsst.sims.maf.utils.opsimUtils import *
        from opsimUtils import *
        import matplotlib.pyplot as plt
        import pandas as pd
        import lsst.sims.maf.slicers as slicers
        import lsst.sims.maf.metrics as metrics
        import lsst.sims.maf.metricBundles as metricBundles
        import lsst.sims.maf.db as db
        import lsst.sims.maf.plots as plots
        from lsst.sims.maf.metrics import BaseMetric
        from lsst.sims.maf.utils import m52snr, sigma_slope, astrom_precis
        from builtins import zip
```

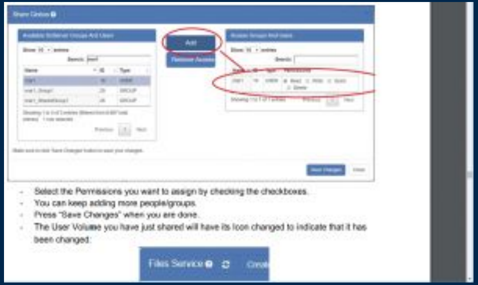
UPDATE 2020-08-06 10am EDT: or, you can add lines like the following to the import cell in the notebook:



```
# add opsimUtils module path to search
import sys
sys.path.insert(0, '../Scripts_NBs/')
from opsimUtils import *
```

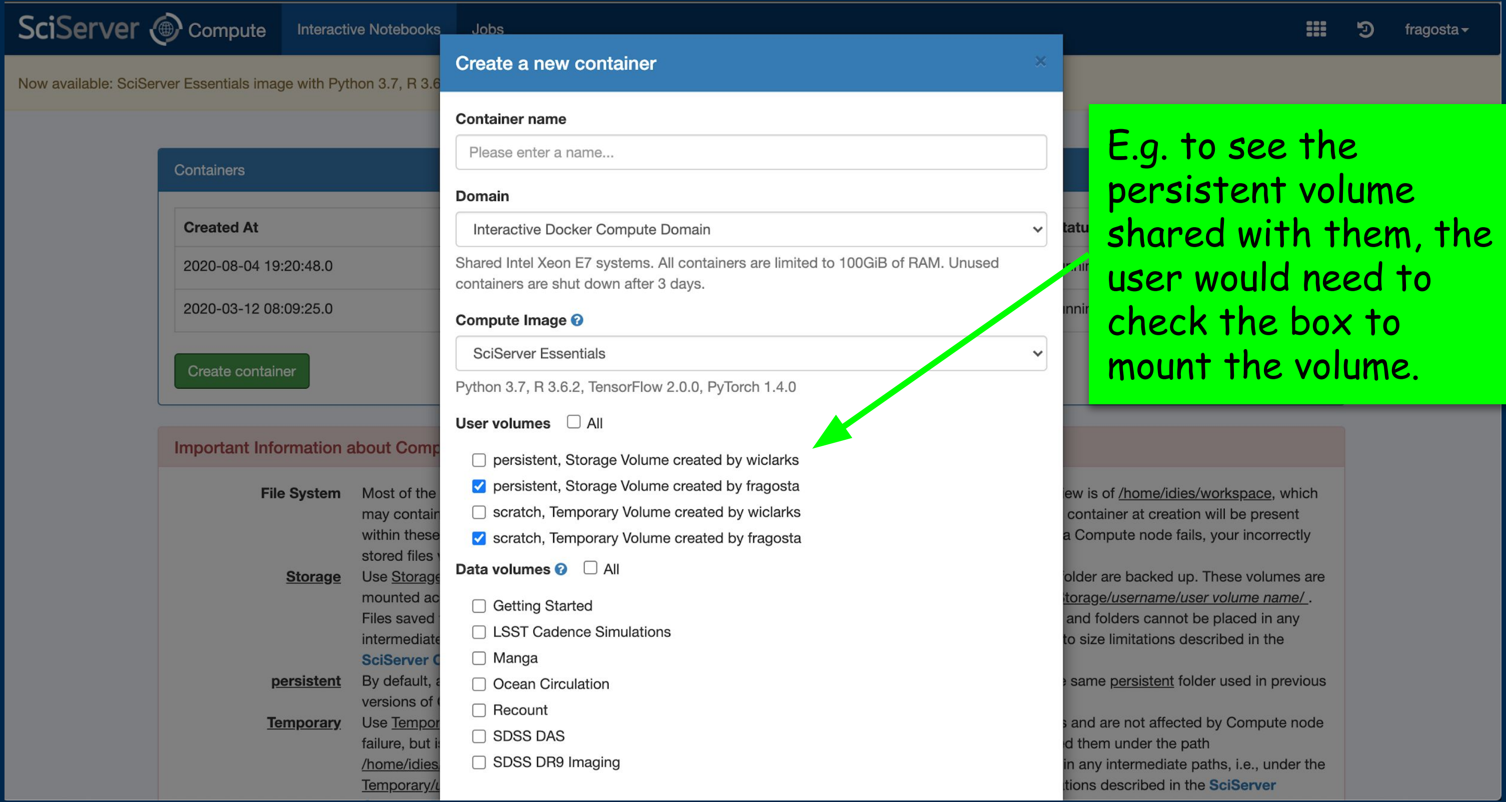
Learning from experience: sharing MAF output on Sciserver

- Handy if, say, a colleague has a folder with 76 sets of hard-won MAF output in their workspace that you want to use as input to aggregate the FoMs!
- User volumes can be shared with other users. See here:
    - http://www.sciserver.org/support/how-to-use-sciserver/
        - and click “File Management” and then “How do I share a User Volume?”
    - Permission is given user-by-user, i.e. the person doing the sharing has to specify which users are granted access and the type of access
- To view the shared folder in jupyter or the terminal, the person you shared the folder with, needs to mount the folder at the stage of creation of the container (see next slide). [*thanks, Weixiang!*]



- Select the Permissions you want to assign by checking the checkboxes.
- You can keep adding more people/groups.
- Press "Save Changes" when you are done.
- The User Volume you have just shared will have its Icon changed to indicate that it has been changed:

Learning from experience: accessing a shared User Volume



Now available: SciServer Essentials image with Python 3.7, R 3.6

Containers

Created At
2020-08-04 19:20:48.0
2020-03-12 08:09:25.0

Create container

Important Information about Comp

File System Most of the ... may contain ... within these ... stored files ...

Storage Use Storage ... mounted ac... Files saved ... intermediate ... SciServer ...

persistent By default, ... versions of ...

Temporary Use Tempor... failure, but i... /home/idies/... Temporary/...

...ew is of /home/idies/workspace, which ... container at creation will be present ... a Compute node fails, your incorrectly

...older are backed up. These volumes are ...torage/username/user volume name/ . ... and folders cannot be placed in any ... to size limitations described in the

... same persistent folder used in previous

... and are not affected by Compute node ... d them under the path ... in any intermediate paths, i.e., under the ... tions described in the SciServer

Create a new container

Container name

Please enter a name...

Domain

Interactive Docker Compute Domain

Shared Intel Xeon E7 systems. All containers are limited to 100GiB of RAM. Unused containers are shut down after 3 days.

Compute Image

SciServer Essentials

Python 3.7, R 3.6.2, TensorFlow 2.0.0, PyTorch 1.4.0

User volumes □ All

- □ persistent, Storage Volume created by wiclarks
- ☑ persistent, Storage Volume created by fragosta
- □ scratch, Temporary Volume created by wiclarks
- ☑ scratch, Temporary Volume created by fragosta

Data volumes □ All

- □ Getting Started
- □ LSST Cadence Simulations
- □ Manga
- □ Ocean Circulation
- □ Recount
- □ SDSS DAS
- □ SDSS DR9 Imaging

E.g. to see the persistent volume shared with them, the user would need to check the box to mount the volume.

[Demo on sciserver running the notebook]

Might show this with byFamily=False

https://github.com/RichardsGroup/LSST_OpSim/blob/master/contrib/05_metricToFOM.ipynb